DEPOSITOR
4 OCT 1978

NAVAL RESEARCH LOGISTICS QUARTERLY

OCT 20 1978
CALIF 93940

JUNE 1978
VOL. 25, NO. 2

OFFICE OF NAVAL RESEARCH

NAVSO P-1278

NAVAL RESEARCH LOGISTICS QUARTERLY

EDITORIAL BOARD

Marvin Denicoff, *Office of Naval Research*, Chairman

Murray A. Geisler, *Logistics Management Institute*

W. H. Marlow, *The George Washington University*

Bruce J. McDonald, *Office of Naval Research Tokyo*

Ex Officio Members

Thomas C. Varley, *Office of Naval Research*
Program Director

Seymour M. Selig, *Office of Naval Research*
Managing Editor

MANAGING EDITOR

Seymour M. Selig
Office of Naval Research
Arlington, Virginia 22217

ASSOCIATE EDITORS

Frank M. Bass, *Purdue University*
Jack Borsting, *Naval Postgraduate School*
Leon Cooper, *Southern Methodist University*
Eric Denardo, *Yale University*
Marco Fiorello, *Logistics Management Institute*
Saul I. Gass, *University of Maryland*
Neal D. Glassman, *Office of Naval Research*
Paul Gray, *University of Southern California*
Carl M. Harris, *Mathematica, Inc.*
Arnoldo Hax, *Massachusetts Institute of Technology*
Alan J. Hoffman, *IBM Corporation*
Uday S. Karmarkar, *University of Chicago*
Paul R. Kleindorfer, *University of Pennsylvania*
Darwin Klingman, *University of Texas, Austin*
Kenneth O. Kortanek, *Carnegie-Mellon University*
Charles Kriebel, *Carnegie-Mellon University*
Jack Laderman, *Bronx, New York*
Gerald J. Lieberman, *Stanford University*
Clifford Marshall, *Polytechnic Institute of New York*
John A. Muckstadt, *Cornell University*
William P. Pierskalla, *Northwestern University*
Thomas L. Saaty, *University of Pennsylvania*
Henry Solomon, *The George Washington University*
Wlodzimierz Szwarc, *University of Wisconsin, Milwaukee*
James G. Taylor, *Naval Postgraduate School*
Harvey M. Wagner, *The University of North Carolina*
John W. Wingate, *Naval Surface Weapons Center, White Oak*
Shelemyahu Zacks, *Case Western Reserve University*

The Naval Research Logistics Quarterly is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Information for Contributors is indicated on inside back cover.

The Naval Research Logistics Quarterly is published by the Office of Naval Research in the months of March, June, September, and December and can be purchased from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Subscription Price: $11.15 a year in the U.S. and Canada, $13.95 elsewhere. Cost of individual issues may be obtained from the Superintendent of Documents.

The views and opinions expressed in this Journal are those of the authors and not necessarily those of the Office of Naval Research.

Issuance of this periodical approved in accordance with Department of the Navy Publications and Printing Regulations, P-35 (Revised 1-74).

DETERMINING CONFIDENCE BOUNDS FOR HIGHLY RELIABLE COHERENT SYSTEMS BASED ON A PAUCITY OF COMPONENT FAILURES*

Janet M. Myhre and Andrew M. Rosenfeld

Claremont Men's College
Claremont, California

Sam C. Saunders

Washington State University
Pullman, Washington

ABSTRACT

A computationally simple method for obtaining confidence bounds for highly reliable coherent systems, based on component tests which experience few or no failures, is given. Binomial and Type I censored exponential failure data are considered. Here unknown component unreliabilities are ordered by weighting factors, which are firstly presumed known then sensitivity of the confidence bounds to these assumed weights is examined and shown to be low.

1. INTRODUCTION

Previously, confidence bounds for general coherent structures have been obtained by using asymptotic methods, such as likelihood Ratio [6], Maximum Likelihood [8], or Modified Maximum Likelihood [1], by using Bayesian methods [7], or by assuming equal reliabilities for all components. Asymptotic methods may be inaccurate at higher percentiles unless the number of failures is very large. With Bayesian methods the possibility of inadvertently influencing a decision through the selection of a prior distribution, when the number of failures is small, is well known. Finally, because the assumption of equal reliabilities of the components may not always be fulfilled, the accuracy of a bound obtained by using this assumption could be in doubt. What we propose here is to use engineering knowledge, which can often be gained from accelerated life tests, material qualification tests, or laboratory tests of components, that provides information about the parameter space. It is felt that this intermediate ground avoids some of the objections raised by Bayesians concerning "classical" statisticians who operate under an assumption of total ignorance about the parameter space. Moreover, it attempts to avoid the subjectivity which often seems to hinder the acceptance of Bayesian methods.

For special structures, such as series structures (and in some cases parallel and series-parallel structures), exact methods [9] and additional approximate and asymptotic methods

*Research supported in part by the Office of Naval Research Contract N00014-76-C-0695.

[4,5,9] for obtaining system confidence bounds have been developed. The accuracy of these approximate bounds has been studied in specific sample cases for structures of order two or three [5,9]. In this paper some comparisons are made between the bounds obtained by the weighting method developed here and approximate (asymptotic) bounds for special structures where such bounds can be calculated.

2. BINOMIAL COMPONENT FAILURE DISTRIBUTIONS

Since the unreliability of any practical system must be low, that of any component must be even lower. The commonly used technique of obtaining confidence bounds for the probability of success from sequences of Bernoulli trials will not be applicable here, because virtually all of the components will have experienced no failures during their acceptance testing. Extending an idea utilized by Lomnicki [3], we will examine the probability of system failures expressed in terms of the least reliable component. The estimate of this quantity is then used to construct a lower confidence bound on the system reliability.

A qualification test for each component consists of a number of Bernoulli trials of nominally identical components. Given there are n_i trials with X_i failures for the i^{th} component, then it is assumed that the number of failures has a binomial distribution where q_i is the unreliability and n_i is the number of observations. We denote this by

$$X_i \sim \beta(n_i, q_i) \qquad \text{for } i = 1, 2, \ldots, m.$$

Assume that q_i may be expressed for each i by

$$q_i = a_i q \text{ where } q = \max_{i=1}^{m} q_i \text{ and } 0 < a_i \leqslant 1. \tag{2.1}$$

For the present assume that the a_i are known *a priori*. In practice we may have obtained the a_i from reliability goals and/or prediction reliabilities. In order to obtain a confidence bound for q, we proceed as follows: Since the q_i are small the distribution of X_i may be accurately approximated by Poisson distribution with mean $\lambda_i = n_i q_i$. Thus, assume

$$X_i \sim P(\lambda_i) \equiv P(n_i q_i). \tag{2.2}$$

(For q_i as large as .01, this approximation is valid for n_i is small as 10.) An upper confidence bound for q is obtained in the usual way since

$$\sum_{i=1}^{m} X_i \sim P\left(\sum_{i=1}^{m} \lambda_i\right)$$

where

$$\lambda \equiv \sum_{i=1}^{m} \lambda_i = \sum_{i=1}^{m} n_i q_i = q \cdot \sum_{i=1}^{m} a_i n_i.$$

The $100\beta\%$ upper confidence bound for λ, call it λ_u, is the value of b for which

$$\sum_{j=0}^{k} e^{-b} b^j / j! = 1 - \beta, \text{ where } k = \sum_{i=1}^{m} X_i.$$

It follows that the $100\beta\%$ upper bound for the unreliability q, call it q_u, is

$$q_u = \lambda_u / \sum_{i=1}^{m} a_i n_i. \tag{2.3}$$

We point out that the Poisson approximation to the Binomial as used above is *not necessary* for the calculation of this type of bound; however, it does greatly facilitate the computation.

In order to obtain a lower confidence bound for the reliability of a coherent system with component reliabilities $\underline{p} = (p_1,\ldots,p_m)$, let

$$p_i = 1 - q_i = 1 - a_i q \text{ for } i = 1, \ldots, m$$

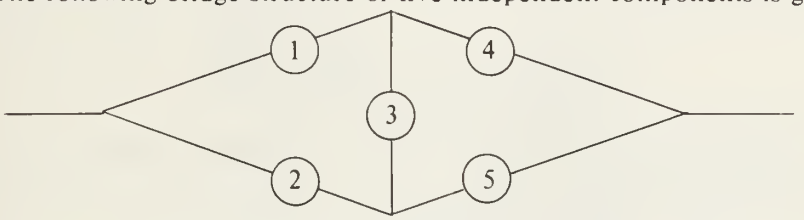
where a_i and q are defined as in equation (2.1). System reliability, $h(\underline{p})$, may now be expressed as a function of q alone. Let the induced function be denoted by the equation

$$h(p) = h(1 - a_1 q, \ldots, 1 - a_m q) = k(q;\underline{a})$$

where $\underline{a} = (a_1,\ldots,a_m)$. The function $k(q;\underline{a})$ is strictly decreasing as a function of q. Hence a lower confidence bound on reliability, $k(q,\underline{a})$, is $k(q_u;\underline{a})$ where q_u is an upper confidence bound on q.

To illustrate these concepts, consider the following examples.

Example 2.1: The following bridge structure of five independent components is given:



If the component reliabilities are p_i for $i=1, \ldots, 5$ then the system reliability is given by

$$h(\underline{p}) = p_1 p_4 + p_2 p_5 + p_1 p_3 p_5 + p_2 p_3 p_4$$
$$- p_1 p_2 p_3 p_4 - p_1 p_2 p_3 p_5 - p_1 p_2 p_4 p_5 - p_1 p_3 p_4 p_5$$
$$- p_2 p_3 p_4 p_5 + 2p_1 p_2 p_3 p_4 p_5.$$

Rewriting in terms of the unreliabilities $1 - p_i = q_i = a_i q$ for $i = 1, \ldots, 5$ yields

$$k(q,\underline{a}) = 1 - q^2(a_1 a_2 + a_4 a_5) - q^3(a_1 a_3 a_5 + a_2 a_3 a_4)$$
$$+ q^4(a_1 a_2 a_3 a_4 + a_1 a_2 a_3 a_5 + a_1 a_2 a_4 a_5 + a_1 a_3 a_4 a_5 + a_2 a_3 a_4 a_5)$$
$$- 2q^5(a_1 a_2 a_3 a_4 a_5).$$

From engineering analyses it is known that components 1, 2, 4 and 5 have the same unreliability. However, it is also assumed that component 3 is only 3/10 as unreliable as the other components. Assume the following weights and test results.

component	a_i	n_i	x_i
1	1	10	0
2	1	10	0
3	.3	20	0
4	1	10	0
5	1	10	0

Suppose a 90% confidence bound is desired. Since $k = \sum_{i=1}^{5} x_i = 0$, λ_u is the value of b for which $e^{-b} = .10$, so $\lambda_u = 2.303$. Hence an upper bound on the unreliability q, at the 90%

level, is given by $q_u = \lambda_u / \sum_{i=1}^{5} a_i n_i = 2.303/46 = 0.050$. Finally, the desired 90% lower confidence bound on reliability is given by $h(q_u, a) = 0.995$.

For such bridge structures it is not possible to compute either Approximately Optimum [4] or Poisson Approximation [9] bounds. Moreover, asymptotic methods are generally not applicable unless failures are observed. A comment is now in order concerning an empirical relationship between the a_i and n_i. For industrial problems we have often found that sample sizes generally are not equal but are roughly proportional to the unreliabilities with the *most unreliable component having the smallest sample size.* One reason for this may be that specialized, complex equipment often tends to be both unreliable and expensive to test.

Example 2.2: Assume a series structure of order five has known weighting factors a_i and sample sizes n_i where $i = 1,2, \ldots, 5$.

component	a_i	n_i
1	1/2	40
2	1	20
3	1/4	80
4	1/2	40
5	1/2	40

Assuming one failure on component 2, $\lambda_u = 3.89$ at the 90% confidence level. Using (2.3) we find that $q_u = .039$. Since $h(q;\underline{a}) = \prod_{i=1}^{5} (1 - a_i q)$, our confidence bound on system reliability is

$$h(q_u;\underline{a}) = .897.$$

The bounds obtained by approximate or asymptotic methods [4,5,9] are much lower than this bound, for example

Approximately Optimum (AO) Bound = .820
Modified Maximum Likelihood Bound (MML1) = .819
Poisson Approximation (PA) Bound = .806

Of course if one assumes more, one should conclude more so these particular differences may not be compelling.

3. SENSITIVITY OF CONFIDENCE BOUNDS TO ASSUMED WEIGHTS

The question that arises is, what is the real difference between the bounds obtained presuming that $\underline{a}$ is known when in fact it may not be. A measure of the error caused by this supposition upon the bound obtained should be found. Let the estimates made up by the experimenter for the values of $\underline{a} = (a_1, \ldots, a_m)$ be denoted by $\underline{\alpha} = (\alpha_1, \ldots, \alpha_m)$. The estimate of the upper bound constructed using $\underline{\alpha}$ in equation (2.3) will be denoted by

$$\tilde{q}_u = \lambda_u / \sum_{i=1}^{m} \alpha_i n_i. \tag{3.1}$$

Example 3.1: Differences between the exact bounds obtained in example 2.2 and bounds obtained using various $\underline{\alpha}$ are given below. The corresponding AO, MMLI and PA bounds are also given. (The asterisk by the a_i or α_i indicates the component on which the failure is assumed to have occurred.)

Component	n_i	a_i	$\alpha_i^{(1)}$	$\alpha_i^{(2)}$	$\alpha_i^{(3)}$	$\alpha_i^{(4)}$	$\alpha_i^{(5)}$
1	40	1/2	1	1/2	1	1/5	1/100
2	20	1*	1/2	1/4	1	1*	1*
3	80	1/4	1/4	1*	1*	1/10	1/100
4	40	1/2	1*	1/2	1	1/5	1/100
5	40	1/2	1	1/2	1	1/5	1/100

Bounds

Weighting Method	.897	.906	.928	.915	.878	.817
AO	.820	.871	.882	.882	.820	.820
MMLI	.819	.916	.950	.950	.819	.819
PA	.806	.806	.806	.806	.806	.806

Note the weighting in case 5 which must be assumed in order to obtain confidence bounds which are close to the AO or MMLI methods. In general, we would not expect the engineering estimates of the ratios of the unreliabilities of the components to be off by a factor of 25 from the correct weights for a series structure.

We now introduce the following notation for subsequent use:

$$\bar{a}_i = a_i / \sum_{i=1}^{m} a_i \qquad n_{\bar{a}} = \sum_{i=1}^{m} \bar{a}_i n_i = \Sigma a_i n_i / \Sigma a_i$$

$$\bar{\alpha}_i = \alpha_i / \sum_{i=1}^{m} \alpha_i \qquad n_{\bar{\alpha}} = \sum_{i=1}^{m} \bar{\alpha}_i n_i = \Sigma \alpha_i n_i / \Sigma \alpha_i$$

$$n_{(1)} = \min(n_1, \ldots, n_m)$$

THEOREM 3.1: Let $h(q_u;a)$ denote the true lower confidence bound for a series system of order m, and let $h(\tilde{q}_u;\alpha)$ be the estimated lower bound. Then the absolute difference, $\mathfrak{D}$, between the true and estimated confidence bounds for a series system of binomial components of order m satisfies:

$$\mathfrak{D} = |h(q_u;\underline{a}) - h(\tilde{q}_u;\underline{\alpha})| < \frac{s - s^{m+1}}{1-s} \text{ where } s = \frac{\lambda_u}{n_{(1)}} \tag{3.2}$$

and if

$$n_{\bar{a}} = n_{\bar{\alpha}} \equiv \bar{n} \tag{3.3}$$

then

$$|h(q_u;\underline{a}) - h(\tilde{q}_u;\underline{\alpha})| < \frac{t^2}{2}\left[1 - \frac{1}{m}\right] + \frac{t^3}{1-t}\left[1 - t^{m-2}\right] \text{ where } t = \frac{\lambda_u}{\bar{n}} \tag{3.4}$$

PROOF: By definition we obtain for the reliability of a series system

$$\prod_{i=1}^{m} h(p) = \prod_{i=1}^{m} h(1 - a_i q) = 1 - s_{1a} q + s_{2a} q^2 \pm \ldots \pm s_{ma} q^m \tag{3.5}$$

where

$$s_{1a} = \sum_{i=1}^{m} a_i \quad s_{2a} = \sum_{i<j}^{m} a_i a_j \quad \ldots \quad s_{ma} = \prod_{i=1}^{m} a_i.$$

From (3.5),

$$\mathfrak{D} \leqslant |s_{1a}q_a - s_{1\alpha}\tilde{q}_\alpha| + |s_{2a}q_a - s_{2\alpha}\tilde{q}_\alpha| + \ldots + |s_{ma}q_a - s_{m\alpha}\tilde{q}_\alpha|. \tag{3.6}$$

Recalling (2.3) and (3.1), (3.6) becomes

$$\mathfrak{D} \leqslant \lambda_u \left| \frac{\Sigma a_i}{\Sigma a_i n_i} - \frac{\Sigma \alpha_i}{\Sigma \alpha_i n_i} \right| + \lambda_u^2 \left| \frac{\sum_{i<j} a_i a_j}{(\Sigma a_i n_i)^2} - \frac{\sum_{i<j} \alpha_i \alpha_j}{(\Sigma \alpha_i n_i)^2} \right| + \ldots . \tag{3.7}$$

$$\ldots + \lambda_u^m \left| \frac{\Pi a_i}{(\Sigma a_i n_i)^m} - \frac{\Pi \alpha_i}{(\Sigma \alpha_i n_i)^m} \right|.$$

To establish (3.2), we note that if $a_i > 0$ for all i, then

$$\frac{\sum_{i_1 < \quad < i_k} a_{i_1} a_{i_2} \ldots a_{i_k}}{(\Sigma a_i n_i)^k} \leqslant \left(\frac{1}{n_{(1)}} \right)^k. \tag{3.8}$$

By (3.8) and the fact that for positive x and y, $|x-y| \leqslant \max(x,y)$, expression (3.7) may be bounded above by

$$\sum_{k=1}^{m} \left(\frac{\lambda_u}{n_{(1)}} \right)^k = \frac{s - s^{m+1}}{1-s} \text{ where } s = \frac{\lambda_u}{n_{(1)}}.$$

Assuming (3.3) we note that (3.7) becomes

$$\mathfrak{D} \leqslant \left(\frac{\lambda_u}{\bar{n}} \right)^2 \left| \sum_{i<j} \bar{a}_i \bar{a}_j - \sum_{i<j} \bar{\alpha}_i \alpha_j \right| + \ldots + \left(\frac{\lambda_u}{\bar{n}} \right)^m \left| \prod_1^m \bar{a}_i - \prod_1^m \bar{\alpha}_i \right|. \tag{3.9}$$

Let $t = \lambda_u / \bar{n}$, then (3.9) becomes

$$\mathfrak{D} \leqslant t^2 \left| \sum_{i<j} \bar{a}_i \bar{a}_j - \sum_{i<J} \bar{\alpha}_i \bar{\alpha}_j \right| + \ldots + t^m |\Pi \bar{a}_i - \Pi \bar{\alpha}_i|. \tag{3.10}$$

Using the method of La Grange multipliers on the first term of (3.10), the maximum value of $\sum_{i<J} \bar{a}_i \bar{a}_j$, subject to the restriction that $\sum_{i=1}^{m} \bar{a}_i = 1$, is obtained at $a_j = 1/m$ for $j = 1, 2, \ldots, m$. Thus

$$\left| \sum_{i<j} \bar{a}_i \bar{a}_j - \sum_{i<j} \bar{\alpha}_i \bar{\alpha}_j \right| \leqslant \max \left(\sum_{i<j} \bar{a}_i \bar{a}_j, \sum_{i<j} \bar{\alpha}_i \bar{\alpha}_j \right) \leqslant \frac{m-1}{2m}. \tag{3.11}$$

The other differences between the corresponding symmetric polynomials are certainly less than unity. Therefore, assuming them to be unity and performing the geometric sum, $\sum_{i=3}^{m} t^i = (t^3 - t^{m+1})/(1-t)$, we obtain a bound on the remaining terms. Assuming (3.3), we obtain

$$D \leqslant t^2 \left\{ \frac{n-1}{2n} + \frac{t^3 - t^{m+1}}{1-t} \right. \quad \text{thus establishing (3.4).} \qquad \text{Q.E.D.}$$

For small $n_{(1)}$, the bound obtained from (3.2) may be too large, particularly if one or more failures have been observed. For example the absolute error bound obtained for the data in Examples 2.2 and 3.1 is .24. In practice, however, when a sampling scheme such as that given in Example 2.2 is used, there is adequate engineering knowledge behind the choice of the

α_i so that errors of such large magnitude are not encountered. If little is known about the weighting then equal component sample sizes are recommended.

Example 3.2: Consider a series system with equal sample sizes for each component:

component	a_i	n_i	α_i
1	1	20	1/2
2	1/100	20	1
3	1/100	20	1/4

Assume no failures, then for a 90% confidence level $\lambda_u = 2.3026$.

The actual error in computing the series system confidence bound is

$$|h(\tilde{q}_u;\underline{\alpha}) - h(q_u;\underline{a})| = |.889 - .885| = .004$$

For this example, the AO bound is .892, a difference of .007 from the correct bound. However, by Theorem 3.1 an absolute bound on the error for the weighting factor method would be

$$t^2\left(\frac{m-1}{2m}\right) + \frac{t^3 - t^{m+1}}{1-t} = .0059$$

where

$$t = \frac{\lambda_u}{\bar{n}} = \frac{2.3026}{20} \text{ and } m = 3.$$

Obviously if the true bound were higher, say .887, then the AO would differ from the true by only .005 and the actual error would be only .002. However, the point to be made is that we often have more information available than simply the structure and the sample size and when we do, it should be used.

Also it should be noted that if the order m of the structure were increased to 20 the absolute error bound would still be only .0080.

We now wish to extend this error determination to a parallel structure of m components. In general, we denote the reliability of a parallel structure of binomial components by:

$$h(\underline{p}) = 1 - \prod_{i=1}^{m}(1-p_i).$$

Writing $h(p)$ in terms of q we obtain

$$h(p) = h(q;\underline{a}) = 1 - \prod_{i=1}^{m} a_i q.$$

THEOREM 3.2: Let $h(q_u;\underline{a})$ denote the true lower confidence bound for $h(\underline{p})$, and let $h(\tilde{q}_u;\underline{\alpha})$ denote the estimated lower confidence bound for $h(\underline{p})$. We obtain

$$\mathfrak{D} = |h(q_u;\underline{a}) - h(\tilde{q}_u;\underline{\alpha})| \leqslant \frac{(\lambda_u/m)^m}{\prod_{i=1}^{m} n_i}. \tag{3.12}$$

PROOF: (3.12) is proved by noting that

$$\mathcal{D} = \left| \prod_{i=1}^{m} a_i q_u - \prod_{i=1}^{m} \alpha_i \tilde{q}_u \right| = (\lambda_u)^m \left| \frac{\prod_{i=1}^{m} a_i}{(\Sigma a_i n_i)^m} - \frac{\prod_{i=1}^{m} \alpha_i}{(\Sigma \alpha_i n_i)^m} \right|. \tag{3.13}$$

For $a_i > 0$, the quantity

$$\mathcal{P} = \frac{\prod_{i=1}^{m} a_i}{(\Sigma a_i n_i)^m} = \exp \left[\ln \frac{\prod_{i=1}^{m} a_i}{(\Sigma a_i n_i)^m} \right]$$

is maximized when $\Sigma \ln a_i - m \ln (\Sigma a_i n_i)$ is maximized. This maximum occurs at $a_j = \left(\sum_{i=1}^{m} a_i n_i \right) / m n_j$ for $j = 1, 2, \ldots, m$. Thus the maximum value of $\mathcal{P}$ is $1/(m^m \Pi n_i)$. Since each term in the difference

$$\left| \frac{\Pi a_i}{(\Sigma a_i n_i)^m} - \frac{\Pi \alpha_i}{(\Sigma \alpha_i n_i)^m} \right|$$

is positive, the absolute value of the difference cannot exceed $1/(m^m \Pi n_i)$. From this (3.13) is established:

$$D \leqslant \frac{(\lambda_u / m)^m}{\prod_{i=1}^{m} n_i}. \qquad \text{Q.E.D.}$$

EXAMPLE 3.3: Consider a parallel structure of order two where no failures occur in testing

component	a_i	n_i	α_i
1	1/2	40	1
2	1	20	1/10

For a 90% confidence bound $\lambda_u = 2.3026$. The 90% lower confidence bounds for system reliability are:

$$h(q_u; \underline{a}) = 1 - q_u^2 \prod_{i=1}^{2} a_i = .99834$$

$$h(\tilde{q}_u; \underline{\alpha}) = 1 - \tilde{q}_u^2 \prod_{i=1}^{2} \alpha_i = .99970$$

The actual difference between these bounds is 1.36×10^{-3}. From Theorem 3.2 the maximum possible error due to incorrectly choosing the weights is 1.66×10^{-3}. But in this instance the AO bound is .99829 with a difference from the correct bound of 0.5×10^{-4}.

It should be noted that the weighting method bounds are easier to compute than the AO bounds.

We now consider a series parallel structure of order $m = \sum_{i=1}^{k} m_i$ which consists of k parallel structures in series; the i^{th} parallel structure has m_i components. Partition $\underline{p} = (p_1, p_2, \ldots, p_m)$ into $\underline{p} = (\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_k)$ where $\underline{p}_i = (p_{m_1+\ldots+m_{i-1}+1}, \ldots, p_{m_1+\ldots+m_i})$. The reliability of the system is given by $h(p) = \prod_{i=1}^{k} h_i(\underline{p}_i)$ with $h_i(\underline{p}_i)$ being a parallel structure consisting of m_i components. Now let $q \equiv \max(q_1, \ldots, q_m)$ where again $q_i = a_i q$ for $i = 1, \ldots, m$. The corresponding partition of $\underline{a}$ is $\underline{a} = (\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k)$. Therefore expressing the reliability in terms of q we obtain

$$h(q;\underline{a}) = \prod_{i=1}^{k} h_i(q;\underline{a}_i).$$

Therefore we establish the error on the lower confidence bound estimate for a series parallel system by

THEOREM 3.3: Let $h_i(q_u; \underline{a}_i)$ be the true lower confidence bound for the i^{th} ($i=1,2, \ldots, k$) parallel structure of order m_i, and let $h_i(\tilde{q}_u;\underline{\alpha}_i)$ be the associated estimate of the lower confidence bound. Then an absolute bound on error in estimating the confidence bound for a series parallel system is given by

$$\left| \prod_{i=1}^{k} h_i(q_u;\underline{a}_i) - \prod_{i=1}^{k} h_i(\tilde{q}_u;\underline{\alpha}_i) \right| \leqslant \sum_{i=1}^{k} \epsilon_i \tag{3.14}$$

where

$$\epsilon_i = \left(\frac{\lambda_u}{m_i}\right)^{m_i} \left(\frac{1}{\prod_{j=1}^{m_i} n_j} \right), \tag{3.15}$$

that is ϵ_i is the bound on error of the i^{th} parallel structure. The n_j are the sample sizes of the components made during the qualification tests.

PROOFS: By induction. For the case when $k=1$, see Theorem 3.2.

Let $k=2$. For convenience set

$$f_i = h_i(q_u; \underline{a}_i) \text{ and } \ell_i = h_i(\tilde{q}_u;\underline{\alpha}_i).$$

Then $|f_1 \cdot f_2 - \ell_1 \cdot \ell_2| = |f_1 \cdot f_2 - \ell_1 \cdot \ell_2 - \ell_2 \cdot f_1 + \ell_2 \cdot f_1|$

$$= |f_1(f_2 - \ell_2) + \ell_2(f_1 - \ell_1)| \leqslant |f_1(f_2 - \ell_2)| + |\ell_2(f_1 - \ell_1)|$$

$$\leqslant |f_2 - \ell_2| + |f_1 - \ell_1| = \epsilon_1 + \epsilon_2.$$

Assume expression (3.14) is true for $k=n$. It holds for $k=n+1$ since,

$$\left| \prod_{i=1}^{n+1} f_i - \prod_{i=1}^{n+1} \ell_i \right| = \left| \prod_{i=1}^{n+1} f_i - \prod_{i=1}^{n+1} \ell_i + f_{n+1} \prod_{i=1}^{n} \ell_i - f_{n+1} \prod_{i=1}^{n} \ell_i \right|$$

$$= \left| \prod_{i=1}^{n} \ell_i (f_{n+1} - \ell_{n+1}) + f_{n+1} \left(\prod_{i=1}^{n} f_i - \prod_{i=1}^{n} \ell_i \right) \right|$$

$$\leqslant \left| \prod_{i=1}^{n} \ell_i (f_{n+1} - \ell_{n+1}) \right| + \left| f_{n+1} \left(\prod_{i=1}^{n} f_i - \prod_{i=1}^{n} \ell_i \right) \right|$$

$$\leqslant \left| f_{n+1} - \ell_{n+1} \right| + \left| \prod_{i=1}^{n} f_i - \prod_{i=1}^{n} \ell_i \right| \leqslant \text{(by induction)}$$

$$|f_{n+1} - \ell_{n+1}| + \sum_{i=1}^{n} |f_i - \ell_i| = \sum_{i=1}^{n+1} |f_i - \ell_i| = \sum_{i=1}^{n+1} \epsilon_i. \qquad \text{Q.E.D.}$$

Using a similar partitioning procedure for a parallel series structure of order $m = \sum_{i=1}^{k} m_i$, then this system consists of k series structures in parallel, where the i^{th} series structure has order m_i. Then $\underline{p} = (\underline{\mathbf{p}}_1, \ldots, \underline{p}_k)$, $h(\underline{p}) = 1 - \prod_{j=1}^{k} (1 - h_j(\underline{p}_j))$ and $h(q; \underline{a}) = 1 - \prod_{j=1}^{k} (1 - h_j(q; \underline{a}_j))$. The maximum error made by estimating the lower confidence bound is obtained from

THEOREM 3.4: Let $h(q_u; \underline{a}) = 1 - \prod_{j=1}^{k} (1 - h_j(q_u; \underline{a}_j))$ be the true confidence bound on the reliability of a parallel series structure of order $m = \sum_{i=1}^{k} m_i$. And let $h(\tilde{q}_u; \underline{\alpha}) = 1 - \prod_{j=1}^{k} (1 - h_j(\tilde{q}_u; \underline{\alpha}_j))$ be the associated estimate of this lower confidence bound. Then

$$\mathcal{B} = |h(q_u; \underline{a}) - h(\tilde{q}_u; \underline{\alpha}) \leqslant \sum_{j=1}^{k} \tau_j \tag{3.16}$$

where $\tau_j = \dfrac{s_j - s_j^{m+1}}{1 - s_j}$ and s_j is defined in Theorem 3.1.

PROOF: Expanding $\mathcal{B}$ we obtain

$$\mathcal{B} = \left| \prod_{j=1}^{k} (1 - h_j(q_u; \underline{a}_j)) - \prod_{j=1}^{k} (1 - h_j(\tilde{q}_u; \underline{\alpha}_j)) \right|.$$

By Theorem 3.3,

$$\mathcal{B} \leqslant \sum_{j=1}^{k} |(1 - h_j(q_u; \underline{a}_j)) - (1 - h_j(\tilde{q}_u; \underline{\alpha}_j))|$$

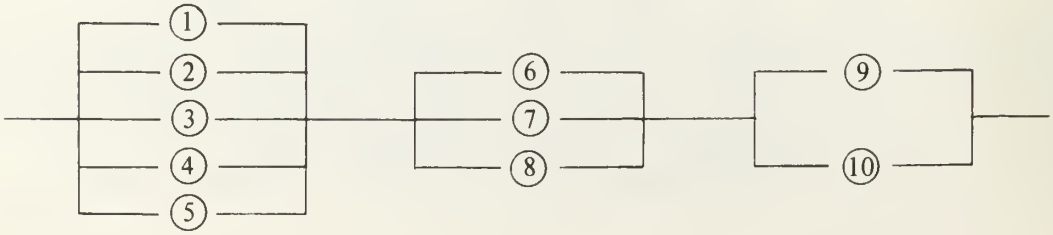
$$= \sum_{j=1}^{k} |h_j(q_u; \underline{a}_j) - h_j(\tilde{q}_u; \underline{\alpha}_j)|$$

where for the i^{th} branch of this parallel structure we define (by Theorem 3.1)

$$\tau_j \equiv |h_j(q_u; \underline{a}_j) - h_j(\tilde{q}_u; \alpha_j)| = \frac{s_j - s_j^{m+1}}{1 - s_j}.$$

$$\text{Thus } \mathcal{B} \leqslant \sum_{j=1}^{k} \tau_j. \qquad \text{Q.E.D.}$$

Example 3.4: Consider the following structure.



component	a_i	n_i	α_i
1	1/4	100	1
2	1/4	100	1
3	1/4	100	1
4	1/4	100	1
5	1/4	100	1
6	1	150	3/4
7	1	150	3/4
8	1	150	3/4
9	3/4	250	1/4
10	3/4	250	1/4

We assume $\Sigma X_i = 1$ and we desire an 80% confidence level. This will yield $\lambda_u = 2.99$. Simple calculations show us

$$\Sigma a_i n_i = 950 \qquad q_u = .00315$$

$$\Sigma \alpha_i n_i = 962.5 \qquad \tilde{q}_u = .00311$$

$$h(q_u; \underline{a}) = (1 - (.25q_u)^5)\ (1 - q_u^3)\ (1 - (.75q_u)^2) = .999994$$

$$h(\tilde{q}_u; \underline{\alpha}) = (1 - \tilde{q}_u^5)(1 - (.75\tilde{q}_u)^3)(1 - (.25\tilde{q}_u)^2) = .999999.$$

We realize an actual error of 5.38×10^{-6}, and based on Theorem 3.3, a maximum possible error of 3.6×10^{-5}.

4. EXPONENTIAL COMPONENT FAILURE DISTRIBUTIONS

We now examine components whose life lengths are known to be exponential. If the qualification test time for the i^{th} component is T_i, where $i=1,2, \ldots, m$, then the number of failures during the interval $[0,T_i]$ follows a Poisson distribution with mean $\lambda_i T_i$. Define

$$\lambda_i = a_i \lambda \text{ where } \lambda = \max_{i=1}^{m} \lambda_i \text{ and } 0 < a_i \leqslant 1. \tag{4.1}$$

Let X_i denote the number of observed failures during $[0,T_i]$, then $X_i \sim P(a_i \lambda T_i)$. If all m components are independent, $\Sigma X_i \sim P(\lambda \Sigma a_i T_i)$.

The upper confidence bound obtained here is completely analogous to the bound obtained in (2.3), i.e., the upper $100\beta\%$ confidence bound on λ, say $\hat{\lambda}$, is

$$\hat{\lambda} = \lambda_u / \Sigma a_i T_i.$$

Reliability at time t, $R(t; \lambda_1, \ldots, \lambda_m)$, may be written as $h(\lambda t; \underline{a})$ where λ and $\underline{a}$ are defined in (4.1). If follows that a lower $100\beta\%$ confidence bound on system reliability is given by $h(\hat{\lambda} t; \underline{a})$ where $\underline{a}$ is the vector of true weights.

5. SENSITIVITY OF CONFIDENCE BOUNDS TO ASSUMED WEIGHTS

When the true weights $\underline{a}$ are estimated by $\underline{\alpha}$ then the estimated $100\beta\%$ lower confidence bound for system reliability becomes

$$h(\tilde{\lambda} t; \underline{\alpha}) \text{ where } \tilde{\lambda} = \lambda_u / \Sigma \alpha_i T_i. \tag{5.1}$$

For the exponential series case we measure error by means of the ratio:

$$\mathcal{R} \equiv \frac{h(\hat{\lambda}t;\, \underline{a})}{h(\tilde{\lambda}t;\underline{\alpha})} \tag{5.2}$$

We will show that in the case of equal component test times for series systems, the above ratio is identically one. That is, the estimate is equal to the true confidence bound for any weighting $\underline{\alpha}$.

THEOREM 5.1: For a series system of order m,

$$\exp -\lambda_u t \left[\frac{1}{T_{(1)}} - \frac{1}{T_{(m)}}\right] \leqslant \mathcal{R} \leqslant \exp \lambda_u t \left[\frac{1}{T_{(1)}} - \frac{1}{T_{(m)}}\right] \tag{5.3}$$

where $T_{(1)} = \min(T_1, \ldots, T_m)$ and $T_{(m)} = \max(T_1, \ldots, T_m)$. If for $i = 1, \ldots, m$, $T_i \equiv T$ then

$$h(\hat{\lambda}t;\, \underline{a}) \equiv h(\tilde{\lambda}t;\underline{\alpha}) = \exp \frac{-\lambda_u t}{T}. \tag{5.4}$$

PROOF: By expanding the ratio in (5.2) we observe

$$\mathcal{R} \equiv \frac{h(\hat{\lambda}t;\, \underline{a})}{h\,\tilde{\lambda}t;\underline{\alpha})} = \frac{\exp \dfrac{-\lambda_u t}{\Sigma a_i T_i}\, \Sigma a_i}{\exp \dfrac{-\lambda_u t}{\Sigma \alpha_i T_i}\, \Sigma \alpha_i} = \exp -\lambda_u t \left|\frac{\Sigma a_i}{\Sigma a_i T_i} - \frac{\Sigma \alpha_i}{\Sigma \alpha_i T_i}\right|.$$

It is simple to see that

$$\frac{1}{T_{(m)}} \leqslant \frac{\Sigma a_i}{\Sigma a_i T_i} \leqslant \frac{1}{T_{(1)}}.$$

Thus

$$-\left[\frac{1}{T_{(m)}} - \frac{1}{T_{(m)}}\right] \leqslant \left|\frac{\Sigma a_i}{\Sigma a_i T_i} - \frac{\Sigma \alpha_i}{\Sigma \alpha_i T_i}\right| \leqslant \left[\frac{1}{T_{(1)}} - \frac{1}{T_{(m)}}\right]. \tag{5.5}$$

Therefore based on (5.5)

$$\exp -\lambda_u t \left[\frac{1}{T_{(1)}} - \frac{1}{T_{(m)}}\right] \leqslant \mathcal{R} \leqslant \exp \lambda_u t \left[\frac{1}{T_{(1)}} - \frac{1}{T_{(m)}}\right].$$

If for all i, $T_i \equiv T$, then the inequality expressed in (5.5) becomes

$$\left|\frac{\Sigma a_i}{\Sigma a_i T_i} - \frac{\Sigma \alpha_i}{\Sigma \alpha_i T_i}\right| \leqslant \left|\frac{1}{T_{(1)}} - \frac{1}{T_{(m)}}\right| \equiv 0$$

therefore the ratio of (5.2) is identically one. That is

$$h(\hat{\lambda}t;\, \underline{a}) = \exp \frac{-\lambda_u t}{T\Sigma a_i}\, \Sigma a_i = \exp \frac{-\lambda_u t}{T\Sigma \alpha_i}\, \alpha_i = h(\tilde{\lambda}t;\underline{\alpha})$$

$$\equiv \exp -\lambda_u t/T. \qquad \text{Q.E.D.}$$

Lieberman and Ross [2] have derived a method for obtaining confidence bounds for series systems whose components have an exponential life distribution. The test statistic used in their

method is based on the sum of simulated system failure times. For Type I censoring the Lieberman-Ross technique will, in general, not utilize all of the test data in the calculation of their confidence bound. In the case of no observed failures the Lieberman-Ross method is not applicable. As shown by the following example, the procedure we propose is not hampered by an absence of qualification test failures.

Example 5.1: Assume that we have twenty components in series and that no failures have been experienced during testing. As is often encountered in practice, the test times are not all equal.

component	a_i	T_i
1	1	10
2 through 20	1/10	100

$\sum_{i=1}^{20} a_i T_i = 200$ and $\sum_{i=1}^{20} a_i = 2.9$. At the 80% confidence level $\lambda_u = 1.61$. Let $t = 1$; then

$$h(\hat{\lambda}t; \underline{a}) = \exp - \left[\frac{1.61}{200}(2.9)\right] = .977.$$

The AO bound is .852. the PA bound is .839.

Assume that the α_i were not chosen correctly but were chosen according to one of the cases given below.

component	$\alpha_i^{(1)}$	$\alpha_i^{(2)}$	$\alpha_i^{(3)}$	$\alpha_i^{(4)}$	T_i
1	1	1	1	1/10	10
2 through 20	1/100	1/2	1	1	100
$h(\hat{\lambda}t;\underline{\alpha})$	.936	.983	.983	.984	

The maximum difference from the true bound is .041. Based on Theorem 5.1 the ratio is bounded by

$$.865 \leqslant \mathcal{R} \leqslant 1.156.$$

Thus we know that regardless of the weights the 80% bound must be greater than .851 (the AO bound is .852). Again, the point to be made here is that we often have more information than simply the structure and the sample size and when we do, it should be used. For this example, using the ratio of Theorem 5.1, it is possible to show that regardless of the weighting the true bound is at least as large as the AO bound (to two significant figures).

In the case of equal test times, say 100, the weighting method bounds are exact and equal to .984. The AO bound is also .984.

THEOREM 5.2: For a parallel system of m exponential components with $0 < \lambda_i t < .1$ for $i = 1, 2, \ldots, m$, the difference

$$D = |h(\hat{\lambda}t;a) - h(\hat{\lambda}t;\underline{\alpha})| \tag{5.6}$$

is approximately bounded above by

$$\frac{(\lambda_u t/m)^m}{\prod_{i=1}^{m} T_i}. \tag{5.7}$$

PROOF: for a parallel system of exponential components

$$\mathcal{R}(t) = 1 - \prod_{i=1}^{m} \left(1 - e^{-\lambda_i t}\right)$$

If $|\lambda t| < .1$ then $\lambda t \simeq 1 - e^{-\lambda t}$. Therefore

$$\mathcal{D} \simeq \left| \prod_{i=1}^{m} \hat{\lambda} t a_i - \prod_{i=1}^{m} \hat{\lambda} t \alpha_i \right| = (\lambda_u t)^m \left| \frac{\prod a_i}{\Sigma a_i T_i)^m} - \frac{\prod \alpha_i}{(\Sigma \alpha_i T_i)^m} \right|.$$

Then by Theorem 3.2, D is approximately less than or equal to

$$\frac{(\lambda_u t/m)^m}{\prod_{i=1}^{m} P_i T_i}$$

Q.E.D.

6. CONCLUSION

The weighting method developed in this paper allows engineering knowledge to be used in a very simple and feasible manner. If little is known about the weighting factors, then we know that for component sample sizes of as small as 20, if there are no failures, the absolute error bound for a series system of infinite order is still less than .008. In practice the actual errors induced due to incorrect choice of the weighting factors are much less than the absolute bounds.

Sensitivity studies show that as qualification test sample sizes (test times) increase, the effect of the weights on the estimated confidence bound decreases. If little is known about the weighting factors, the bound on the maximum possible error induced by different weights may be reduced significantly by imposing equal sample sizes (test times) for the components during testing. Moreover, under the assumption of equivalent component test times in the commonly encountered case of exponential series systems, the confidence bound obtained is exact. The advantages of the weighting metod proposed here lie in the simplicity of the calculations, the applicability to *any* coherent structure when few or no failures occur, the ability to use in an uncomplicated fashion certain types of engineering knowledge to compensate for small sample sizes (test times), and, for larger sample sizes, the demonstrated insensitivity to the choice of weighting factors.

7. ACKNOWLEDGMENT

The authors wish to thank Daniel Giesberg for his extensive work on sensitivity analysis.

REFERENCES

[1] Easterling, Robert G., "Approximate Confidence Limits for System Reliability," Journal of the American Statistical Association **67**, 220-2(March 1972).

[2] Lieberman, Gerald J. and Sheldon M. Ross, "Confidence Intervals for Independent Exponential Series Systems," Journal of the American Statistical Association **66**, 837-40(Dec. 1971).

[3] Lomnicki, Z. A., "Two-Terminal Series-Parallel Networks," Advances in Applied Probability **4**, 109-150(1973).

[4] Mann, Nancy R. and Frank E. Grubbs, "Approximately Optimum Confidence Bounds for System Reliability Based on Component Test Data," Technometrics **16**, 335-47(1974).

[5] Mann, Nancy R., Ray E. Schafer and Nozer D. Singpurwalla, *Methods for Statistical Analysis of Reliability and Life Data* John Wiley and Sons, Inc., New York, 487-524 (1974).

[6] Myhre, J. M. and S. C. Saunders, "Comparison of Two Methods of Obtaining Approximate Confidence Intervals for System Reliability," Technometrics **10**, 37-49(1968).

[7] Raiffa, Howard and Robert Schlaifer, *Applied Statistical Decision Theory*, Division of Research — Harvard Business School, Boston, 1961.

[8] Rao, C. R., *Advanced Statistical Methods in Biometric Research*, John Wiley and Sons, Inc., New York, 1952.

[9] Winterbottom, Alan, "Lower Confidence Limits for Series System Reliability from Binomial Subsystem Data," Journal of the American Statistical Association **69**, 782-787(Sept. 1974).

APPROXIMATIONS TO THE RELIABILITY OF PHASED MISSIONS

Harald Ziehms

Naval Postgraduate School
Monterey, California

ABSTRACT

A system whose configuration (block diagram or fault tree) changes during consecutive time periods (phases) performs a "phased mission." Recently, Esary and Ziehms have shown that any multiphase mission can be transformed into an equivalent, synthetic, single-phase system, and thus that the phased mission problem can in principle be solved by standard reliability methods. We employ these ideas here to study approximations to mission reliability and to develop an algorithm which may be of practical interest. In addition, we extend the reliability calculus of Rubinstein, and Esary and Hayne, based on an approximate hazard transform, to phased missions, and we show how this extended calculus can be used in situations where phases are not of known fixed duration.

1. INTRODUCTION

The technological development of the last two decades, particularly in the areas of space flight, nuclear power generation, and weapons systems, has forced reliability analysts to consider systems whose configurations change over time. "Phased missions" have received attention in the basic papers of Rubin [5] and Weisberg and Schmidt [9], which present computational procedures to approximately predict mission reliability and crew safety for manned space craft, and in the United States Navy reliability manual NAVORD OD 29304 Revision A [8].

Recently, Esary and Ziehms [4] have investigated a phased mission of the following form: A system consists of several independently performing components, each of which functions continuously in time until failure occurs, and remains failed thereafter; repair or replacement is not possible. The system performs a mission which is divided into consecutive time periods, or phases, of known duration. The system configuration, defined as a subset of the components and their functional organization, changes from phase to phase. As is the case with individual components, only two states of the system are recognized, functioning or failed. The mission is successful if the system functions throughout all phases.

Their main result is that any multiphase mission of this type can be transformed into an equivalent, synthetic, single-phase system, and thus that the phased mission problem can be solved in principle by standard reliability methods. They point out, however, that a direct implementation of their transformation could be frustrated by a large number of components in the equivalent system.

*This research was partially supported by the Office of Naval Research (NR 042-300) and the Strategic Systems Project Office (TA 19422).

In this paper we employ the ideas of Esary and Ziehms to study some approximations to mission reliability and to develop an algorithm which may be of practical interest. In addition, we extend the reliability calculus of Rubinstein [6,7] and Esary and Hayne [1], based on an approximate hazard transform, to phased missions, and demonstrate how the extended calculus can be used in situations where phases are not of known fixed duration.

2. PROBLEM FORMULATION AND PREVIOUS RESULTS

Suppose that the system under consideration has n components, labeled $C_1, \ldots, C_n$, with independent times to failure $T_1, \ldots, T_n$. For all times $t \geqslant 0$, define the *performance state indicator vector* of the set of components $\underset{\sim}{X}(t) = (X_1(t), \ldots, X_n(t))$ by $X_k(t) = 1$ *iff* $T_k > t$, and $X_k(t) = 0$ otherwise, $k = 1, \ldots, n$. Assume that the mission is divided into m phases, and that phase j starts at time t_{j-1} and ends at time t_j, $j = 1, \ldots, m$, with $t_0 = 0$. Finally, let ϕ_j be the structure function which describes the configuration (assumed to be coherent) of the system in phase j, $j = 1, \ldots, m$. Then the event that the mission is successful is $\{\phi_1[\underset{\sim}{X}(t_1)] = 1, \ldots, \phi_m[\underset{\sim}{X}(t_m)] = 1\}$, and the mission reliability p can be expressed as

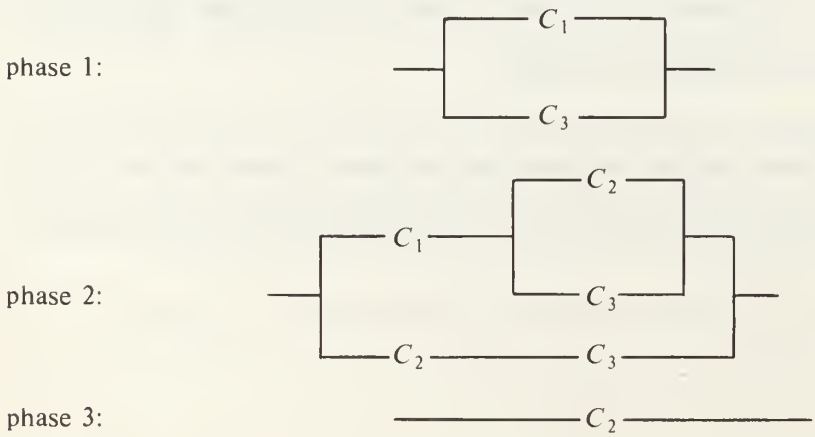
$$p = E\prod_{j=1}^{m} \phi_j[\underset{\sim}{X}(t_j)]. \tag{1}$$

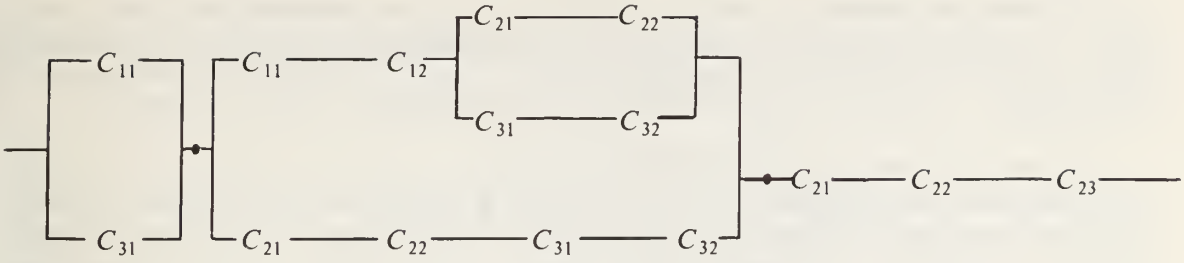
To obtain an equivalent single-phase system, pseudocomponents C_{kj} are introduced whose reliabilities are the conditional phase reliabilities of the original components. Formally, for $k = 1, \ldots, n$ and $j = 1, \ldots, m$, the performance state indicator variable U_{kj} of pseudocomponent C_{kj} has the distribution

$$\begin{aligned} &P[U_{k1} = 1] = P[X_k(t_1) = 1], \\ &P[U_{kj} = 1] = P[X_k(t_j) = 1 \mid X_k(t_{j-1}) = 1], \quad j \neq 1. \end{aligned} \tag{2}$$

The *transformation* is accomplished by replacing, in the configuration for phase j, component C_k by a series system in which the pseudocomponents $C_{k1}, \ldots, C_{kj}$ perform independently with the probabilities of functioning given in (2), and by regarding the transformed phase configurations as subsystems operating in series. As an illustration, consider the following example.

EXAMPLE 1. A system with three components performs a three-phased mission whose phase configurations can be represented by the block diagrams





The reliability of the equivalent system is

$$p = E \prod_{j=1}^{m} \phi_j(\underset{\sim}{U}^{(1)} \underset{\sim}{U}^{(2)} \ldots \underset{\sim}{U}^{(j)}), \tag{3}$$

where $\underset{\sim}{U}^{(j)} = (U_{1j}, \ldots, U_{nj})$ and $U^{(i)} U^{(l)} = (U_{1i}U_{1l}, \ldots, U_{ni}U_{nl})$. The value of p as given by (3) agrees with the value of p as given by (1) ([4], Theorem 3.1), and thus the ordinary reliability of the equivalent system whose components perform independently is the same as the reliability of the original system for its phased mission.

3. SOME BOUNDS ON MISSION RELIABILITY

An obvious first approach to approximating mission reliability—discussed in [4] and repeated here for the sake of completeness—is to compute the reliability of each phase configuration separately and then to multiply the results together. There are at least two choices of component reliabilities to use in doing this: the component *conditional phase reliabilities*

$$\pi_{k1} = P[X_k(t_1) = 1]$$

$$\pi_{kj} = P[X_k(t_j) = 1 | X_k(t_{j-1}) = 1], \qquad j = 2, \ldots, m, \tag{4}$$

which are the reliabilities of the pseudocomponents in the equivalent system, or the component *unconditional reliabilities* through each phase

$$\rho_{kj} = p[X_k(t_j) = 1] = \prod_{i=1}^{j} \pi_{ki}, \qquad j = 1, \ldots, m,$$

$k = 1, \ldots, n$. The first choice leads to approximating mission reliability by

$$\pi_{PRF} = \prod_{j=1}^{m} h_j(\pi_{1j}, \ldots, \pi_{nj}),$$

and the second choice to approximating mission reliability by

$$\rho_{PRF} = \prod_{j=1}^{m} h_j(\rho_{1j}, \ldots, \rho_{nj}), \tag{7}$$

where in both cases h_j, $j = 1, \ldots, m$, are the reliability functions for the phase configurations. (The *reliability function* of a system with structure function ϕ is defined by $h(p_1, \ldots, p_n) = E\phi(X_1 \ldots, X_n)$, where $X_1 \ldots, X_n$ are independent Bernoulli random variables with $P[X_k = 1] = p_k$, $k = 1, \ldots, n$.) The subscript PRF in (6) and (7) is meant to indicate that these approximations are based on phase reliability functions.

It has been shown using (3) ([4], Remark 4.1) that (6) gives an optimistic result and that (7) gives a conservative result; i.e., that for p as given by (1) or (3),

$$\rho_{PRF} \leqslant p \leqslant \pi_{PRF}. \tag{8}$$

The above approximations can be employed only when the reliability functions of all m phases are known. Although to compute them is considerably easier than to compute the overall reliability function for the equivalent system, it may in practical problems still be a formidable task. We will therefore now discuss an approach which avoids these difficulties.

For coherent single-phase systems with independent components, Esary and Proschan [2] have established two bounds on system reliability which do not involve the reliability function: the *minimal path upper bound* and the *minimal cut lower bound.* These bounds, when applied to each phase separately, can be used to approximate mission reliability in the multiphase case. Let h_{UBj} and h_{LBj} denote the minimal path upper bound and the minimal cut lower bound, respectively, for phase configuration j, $j = 1, \ldots, m$. Using basically the same approach as before, and choosing as component reliabilities the conditional phase reliabilities π_{kj} in one case and the (unconditional) reliabilities ρ_{kj} in the other, we obtain the approximations

$$\pi_{PUB} = \prod_{j=1}^{m} h_{UBj}(\pi_{1j}, \ldots, \pi_{nj}) \tag{9}$$

and

$$\rho_{PLB} = \prod_{j=1}^{m} h_{LBj}(\rho_{1j}, \ldots, \rho_{nj}), \tag{10}$$

where the subscripts are to indicate that these approximations are based, respectively, on *phase upper bounds* and *phase lower bounds.* Since the phase configurations are coherent by assumption, $h_{LBj} \leqslant h_j \leqslant h_{UBj}$, $j = 1, \ldots, m$; it thus follows from (6) and (9) that

$$\pi_{PRF} \leqslant \pi_{PUB}, \tag{11}$$

and from (7) and (10) that

$$\rho_{PLB} \leqslant \rho_{PRF}. \tag{12}$$

From (8), (11), and (12) we can conclude that (9) is an upper bound on mission reliability, and (10) is a lower bound on mission reliability.

4. CUT CANCELLATION AND FURTHER BOUNDS

Rubin, Weisberg, and Schmidt used a method to simplify the sequence of phase configurations prior to beginning reliability calculations which has become known as "cut cancellation." Cut cancellation does not affect mission reliability ([4], Remark 4.2) and can be summarized in the following rule:

A minimal cut set in a phase can be cancelled, i.e. omitted from the list of minimal cut sets for that phase, if it contains a minimal cut set of a later phase. The next example illustrates how cut cancellation works.

EXAMPLE 2. Consider the mission of Example 1. The minimal cut sets are

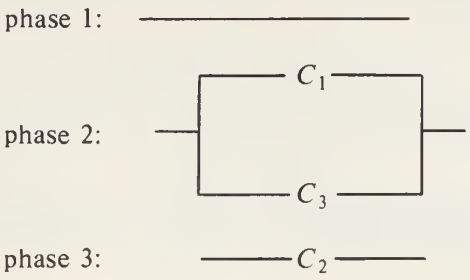
phase 1: $\{C_1, C_3\}$

phase 2: $\{C_1, C_2\}$, $\{C_1, C_3\}$, $\{C_2, C_3\}$

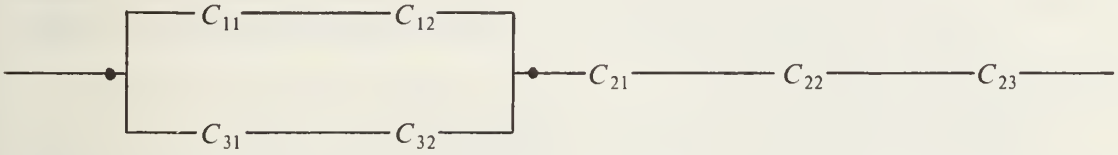
phase 3: $\{C_2\}$.

The phase 1 cut set $\{C_1, C_3\}$ contains the phase 2 cut set $\{C_1, C_3\}$, and thus can be cancelled in phase 1, leaving a configuration which can never fail. Both the phase 2 cut sets $\{C_1, C_2\}$ and

$\{C_2, C_3\}$ contain the phase 3 cut set $\{C_2\}$, so they can be cancelled in phase 2. After cut cancellation, the simplified phase configurations can be represented by the block diagrams



After cancellation, the transformation can be applied to obtain the equivalent system with the block diagram



which is considerably simpler than the equivalent system of Example 1, but has the same reliability.

The methods of approximating mission reliability described in the previous section can also be employed after cut cancellation has been performed. If we denote the reliability functions of the simplified phase configurations by h_j^-, $j = 1, \ldots, m$, the approximations corresponding to π_{PRF} and ρ_{PRF} are

$$\pi_{PRF-CC} = \prod_{j=1}^{m} h_j^-(\pi_{1j}, \ldots, \pi_{nj}) \tag{13}$$

and

$$\rho_{PRF-CC} = \prod_{j=1}^{m} h_j^-(\rho_{1j}, \ldots, \rho_{nj}), \tag{14}$$

respectively, where the added subscript CC indicates that cut cancellation has been performed. Similarly, denoting by h_{UBj}^- and h_{LBj}^- the minimal path upper bound and the minimal cut lower bound, respectively, for the simplified configuration of phase j, $j = 1, \ldots, m$, we obtain the approximations

$$\pi_{PUB-CC} = \prod_{j=1}^{m} h_{UBj}^-(\pi_{1j}, \ldots, \pi_{nj}) \tag{15}$$

and

$$\rho_{PLB-CC} = \prod_{j=1}^{m} h_{LBj}^-(\rho_{1j}, \ldots, \rho n_{nj}). \tag{16}$$

To show that these four approximations are bounds on mission reliability, we observe first that since the simplified phase configurations are coherent, $h_{LBj}^- \leqslant h_j^- \leqslant h_{UBj}^-$, $j = 1, \ldots, m$. It follows from (13) and (15) that

$$\pi_{PRF-CC} \leqslant \pi_{PUB-CC}, \tag{17}$$

(18) $$\rho_{PLB-CC} \leqslant \rho_{PRF-CC}.$$

Further, since the phase reliability functions are not less after cut cancellation than before, i.e. $h_j \leqslant h_j^-$, $j = 1, \ldots, m$, then

(19) $$\pi_{PRF} \leqslant \pi_{PRF-CC}$$

follows from (6) and (13), and

(20) $$\rho_{PRF} \leqslant \rho_{PRF-CC}$$

follows from (7) and (14), where the latter inequality is noted here for further reference only. From (19) and (8) we conclude that π_{PRF-CC} and π_{PUB-CC} are in fact upper bounds on mission reliability.

To establish that ρ_{PRF-CC} and ρ_{PLB-CC} are lower bounds, we need the following remark.

REMARK 1. Let ϕ_j^- be the structure function of the simplified configuration of phase j, $j = 1, \ldots, m$, and let U_{kj}, $k = 1, \ldots, n$, $j = 1, \ldots, m$, be the indicator variables of the pseudocomponents in the equivalent system. Then

$$\prod_{j=1}^{m} E\phi_j^-(\underset{\sim}{U}^{(1)}\underset{\sim}{U}^{(2)} \ldots \underset{\sim}{U}^{(j)}) \leqslant E\prod_{j=1}^{m} \phi_j^-(\underset{\sim}{U}^{(1)}\underset{\sim}{U}^{(2)} \ldots \underset{\sim}{U}^{(j)}).$$

PROOF. The proof uses standard properties of associated random variables which are discussed in Esary, Proschan and Walkup [3].

The simplified phase configurations are coherent, and hence the structure functions ϕ_j^-, $j = 1, \ldots, m$, are nondecreasing. The Bernoulli random variables U_k, $k = 1, \ldots, n$, $j = 1, \ldots, m$, are independent. Therefore $\phi_j^-(\underset{\sim}{U}^{(1)}\underset{\sim}{U}^{(2)} \ldots \underset{\sim}{U}^{(j)})$, $j = 1, \ldots, m$, are associated Bernoulli random variables for which the assertion of the remark holds.

Using (2), (5), and (14) we obtain $\rho_{PRF-CC} = \prod_{j=1}^{m} E\phi_j^-(\underset{\sim}{U}^{(1)}\underset{\sim}{U}^{(2)} \cdot \ldots \underset{\sim}{U}^{(j)})$; since cut cancellation does not affect mission reliability, (3) can be written as $p = E\prod_{j=1}^{m} \phi_j^-(\underset{\sim}{U}^{(1)}\underset{\sim}{U}^{(2)} \ldots \underset{\sim}{U}^{(j)})$. Application of Remark 1 then yields the inequality

(21) $$\rho_{PRF-CC} \leqslant p,$$

which together with (20) establishes the desired results.

5. COMPARISONS OF THE BOUNDS

The magnitudes of the bounds on mission reliability presented in the previous sections, and of the mission reliability itself, can be ordered. This ordering is displayed in Fig. 1 where the superscripts refer to the defining equations and inequalities which are summarized.

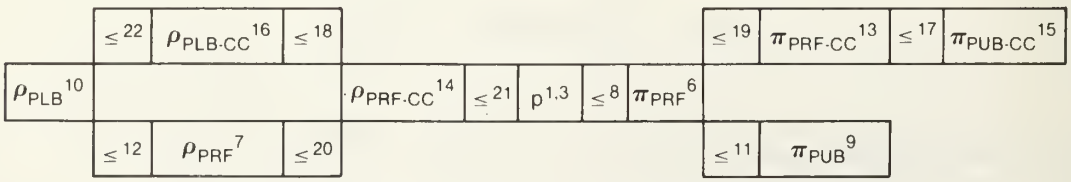


FIGURE 1. Qualitative comparison of the bounds on mission reliability
(The superscripts refer to the defining equations and inequalities.)

No general inequalities can be established between π_{PRF-CC} and π_{PUB}, and between ρ_{PLB-CC} and ρ_{PRF}. In the case of the two upper bounds, cut cancellation on one hand and the use of phase upper bounds instead of phase reliability functions on the other hand both tend to increase the apparent phase reliabilities, the amount of increase depending on the structure of the mission as well as on the comonent reliabilities. In the case of the two lower bounds, ρ_{PLB-CC} tends to be greater than ρ_{PRF} because of cut cancellation, but also smaller because of the use of phase lower bounds instead of phase reliability functions. Again, both the structure of the mission and the component reliabilities determine which of them is greater in a particular case.

The inequality

$$\rho_{PLB} \leqslant \rho_{PLB-CC} \tag{22}$$

has not yet been established formally, but is an obvious consequence of (10), (16), and the fact that $h_{LBj} \leqslant h_{\bar{L}Bj}$, $j = 1, \ldots, m$. A similar inequality between the upper bounds π_{PUB} and π_{PUB-CC}, however, does not exist, because it is not necessarily true that $h_{UBj} \leqslant h_{\bar{U}Bj}$. Since this may not be intuitively obvious, we give the following illustration.

EXAMPLE 3. For the mission of Example 1, the minimal path upper bound for phase 2 is $h_{UB2}(\pi_{12}, \pi_{22}, \pi_{32}) = \pi_{12}\pi_{22} \vee \pi_{12}\pi_{32} \vee \pi_{22}\pi_{32}$ before cut cancellation, and $h_{\bar{U}B2}(\pi_{12}, \pi_{22}, \pi_{32} = \pi_{12} \vee \pi_{32}$ after cut cancellation. Assuming that $\pi_{12} = \pi_{22} = \pi_{32} = \pi$, then $h_{UB2}(\pi) = \pi^2(3-3\pi^2+\pi^4)$ and $h_{\bar{U}B2}(\pi) = \pi(2-\pi)$. For $0 < \pi \leqslant 0.8$, $h_{UB2}(\pi) < h_{\bar{U}B2}(\pi)$, and for $0.9 \leqslant \pi < 1$, $h_{UB2}(\pi) > h_{\bar{U}B2}(\pi)$.

It is also possible to compare the bounds with respect to the computational effort required to compute them. In general, less effort is required to compute the m phase reliability functions separately than to compute one reliability function for the equivalent system; phase bounds are easier to compute than phase reliability functions; and cut cancellation simplifies all reliability calculations, although it requires computational effort itself. The diagram below is an attempt to summarize these observations. Its comparisons may not hold in all cases, but do indicate what is usually true. The symbol $\leftarrow$ stands for "requires less computational effort than."

$$\pi_{PUB-CC} \leftarrow \pi_{PUB} \leftarrow \pi_{PRF-CC} \leftarrow \pi_{PRF}$$

$$\leftarrow p$$

$$\rho_{PLB-CC} \leftarrow \rho_{PLB} \leftarrow \rho_{PRF-CC} \leftarrow \rho_{PRF}$$

6. AN ALGORITHM FOR THE "BEST" LOWER BOUND

Trying to select the best bound from those presented here is a problem whose solution depends on the circumstances of each particular application and cannot be accomplished in general. If one is interested in a conservative rather than an optimistic approximation, and if the system to be analyzed has components with uniformly high conditional reliabilities in all phases, then the qualitative comparisons of the previous section and numerical results suggest that ρ_{PLB-CC} is a good choice. Since these conditions are frequently encountered, an algorithm for computing ρ_{PLB-CC} is given below. Inputs to this algorithm are the phase configurations (in the form of block diagrams, fault trees, structure functions, or complete lists of minimal cut sets or minimal path sets), and estimates of the component conditional phase reliabilities π_{kj}, $k = 1, \ldots, n$, $j = 1, \ldots, m$. If one is willing to assume that components have constant failure rates throughout each phase, then the component conditional phase reliabilities are given by

$$\pi_{kj} = e^{-r_{kj}d_j},$$

where r_{kj} is the *failure rate* of component C_k in phase j, and d_j is the *duration* of phase j, $k = 1, \ldots, n$, $j = 1, \ldots, m$.

ALGORITHM: 1 for computing ρ_{PLB-CC}

(a) For $j = 1, \ldots, m$, find the minimal cut sets for the configuration of phase j.

(b) Perform cut cancellation according to the rule given in Section 4. For $j = 1, \ldots, m$, denote the number of minimal cut sets remaining in phase j by $k(j)$, and the ith minimal cut set remaining in that phase by K_{ji}, $i = 1, \ldots, k(j)$.

(c) For $k = 1, \ldots, n$, compute ρ_{kj} for all $j = 1, \ldots, m$ for which $C_k \epsilon K_{ji}$ for some $i = 1, \ldots, k(j)$, from

$$\rho_{kj} = \prod_{i=1}^{j} \pi_{ki}.$$

(d) Compute ρ_{PLB-CC} from

$$\rho_{PLB-CC} = \prod_{j=1}^{m} \prod_{i=1}^{k(j)} [1 - \prod_{C_k \epsilon K_{ji}} (1 - \rho_{kj})].$$

The notation necessary to formulate this algorithm in precise mathematical terms obscures its basically very simple content. We can restate it in the following more intelligible form:

(a) Find the minimal cut sets for all phase configurations.

(b) Perform cut cancellation.

(c) Compute ρ_{kj} for each phase j in which component C_k is relevant from

$$\rho_{kj} = \prod_{i=1}^{j} \pi_{ki}$$

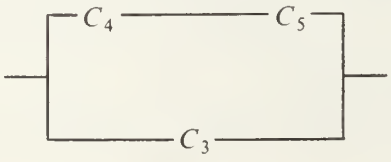
(d) Obtain the "best" lower bound on mission reliability by computing

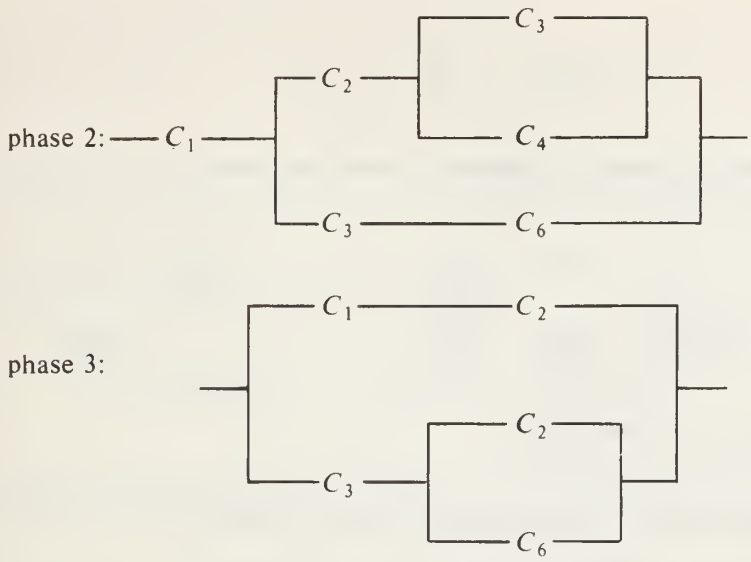
$$\prod_{\{\text{all phases}\}} \quad \prod_{\substack{\{\text{all min cut sets} \\ \text{in each phase}\}}} \quad [1 - \prod_{\substack{\{\text{all components in} \\ \text{each min cut set}\}}} (1-\rho_{kj})]$$

The following example, adapted from Esary and Ziehms [4], illustrates how the algorithm works.

EXAMPLE 4. A system with six components is to perform a three-phased mission. The phase configurations are represented by the block diagrams

phase 1:





The duration of the phases are $d_1 = 30$ min, $d_2 = 2$ hours, and $d_3 = 10$ hours. It is assumed that the components have failure rates r_{kj} which are constant throughout each phase; estimates of their values (in hours^{-1}) are

j/k	1	2	3	4	5	6
1	0.000	0.001	0.020	0.040	0.100	0.000
2	0.020	0.003	0.006	0.010	0.500	0.020
3	0.010	0.002	0.005	0.020	0.500	0.020

A lower bound on mission reliability is wanted.

The application of the algorithm yields the following results:

(a) The minimal cut sets are

phase 1: $\#\{C_3,C_4\}\#$, $\{C_3,C_5\}$

phase 2: $\{C_1\}$, $\#\{C_2,C_3\}\#$, $\#\{C_2,C_6\}\#$, $\{C_3,C_4\}$

phase 3: $\{C_1,C_3\}$, $\{C_2,C_3\}$, $\{C_2,C_6\}$.

(b) The minimal cut sets marked $\#\{\ \}\#$ above are cancelled. The remaining minimal cut sets are

phase 1: $\{C_3,C_5\}$

phase 2: $\{C_1\}$, $\{C_3,C_4\}$

phase 3: $\{C_1,C_3\}$, $\{C_2,C_3\}$, $\{C_2,C_6\}$.

(c) We have to compute ρ_{12}; ρ_{13}; ρ_{23}; ρ_{31}; ρ_{32}; ρ_{33}; ρ_{42}; ρ_{51}; and ρ_{63}. Since in the present case $\pi_{kj} = e^{-r_{kj}d_j}$, we use the equation

$$\rho_{kj} = \prod_{i=1}^{j} \pi_{ki} = \prod_{i=1}^{j} e^{-r_{ki}d_i} = e^{-\sum_{i=1}^{j} r_{ki}d_i}$$

and obtain the following values for ρ_{kj} (rounded to four decimals):

j/k	1	2	3	4	5	6
1			0.9900		0.9512	
2	0.9608		0.9782	0.9608		
3	0.8694	0.9738	0.9305			0.7866

(d) the bound ρ_{PLB-CC} is given by

$$\begin{aligned}\rho_{PLB-CC} &= [1 - (1-\rho_{31})(1-\rho_{51})] \times [1 - (1 - \rho_{12})] \\ &\times [1 - (1-\rho_{32})(1 - \rho_{42})] \times [1 - (1-\rho_{13})(1-\rho_{33})] \\ &\times [1 - (1 - \rho_{23})(1-\rho_{33})] \times [1 - (1-\rho_{23})(1-\rho_{63})].\end{aligned}$$

For the values of ρ_{kj} computed in Step (c), we obtain, rounded to four decimals,

$$\rho_{PLB-CC} = 0.9438.$$

As a comparison, the reliability function for the mission is

$$\begin{aligned}h &= \rho_{12}\rho_{33}(\rho_{23} + \rho_{63} - \rho_{23}\rho_{63}) \\ &+ \rho_{13}\rho_{23}[(1 - \rho_{31})\rho_{42}\rho_{51} + (\rho_{31} - \rho_{32})\rho_{42} + (\rho_{32} - \rho_{33})],\end{aligned}$$

and thus the exact mission reliability, rounded to four decimals, is

$$p = 0.9468.$$

7. AN APPROXIMATE HAZARD TRANSFORM FOR PHASED MISSIONS

Recently, Esary and Hayne [1] extended the scope of a application of a simple reliability calculus of Rubinstein [6, 7] to coherent systems. This calculus uses an approximate hazard transform and leads to conservative approximations to system reliability. We will show here that its scope can be further extended to phased missions.

The hazard transform of a system with reliability function $h(p_1, \ldots, p_n)$ is defined as

$$H(u_1, \ldots, u_n) = -\log h\,(e^{-u_1}, \ldots, e^{-u_n}),$$

where $u_k = -\log p_k$ is the component hazard of component C_k having reliability $p_k, k = 1, \ldots, n$. The approximate hazard transform H^* considered in [1] can be defined by the following rules:

(a) For a system consisting of a single component C_k,

$$H^* = u_k.$$

(b) For a system which is a combination of two modules (subsystems with disjoint sets of components) having approximate hazard transforms H_1^* and H_2^*,

$$H^* = H_1^* + H_2^* \text{ if the combinaton is series}$$
$$H^* = H_1^* H_2^* \quad \text{if the combination is parallel.}$$

(c) For a coherent system with minimal cut sets $K_1, \ldots, K_l$ whose approximate hazard transforms are $H_1^*, \ldots, H_l^*$,

$$H^* = H_1^* + \ \ldots \ + H_l^*.$$

It has been shown ([1], Theorem 2.5) that this approximate hazard transform is conservative, i.e. indicates greater system hazard (less system reliability) than the exact hazard transform.

In the case of a phased mission, we can go one step further and define an *approximate mission hazard transform* by the rule

(d) For a phased mission whose simplified phase configurations have approximate hazard transforms $H_1^*, \ldots, H_m^*$, the approximate mission hazard transform is

$$H^* = H_1^* + \ \ldots \ + H_m^*,$$

where the component hazards are $u_{kj} = -\log \rho_{kj}$, $k = 1, \ldots, n$, $j = 1, \ldots, m$. We will denote the reliability function corresponding to this approximate mission hazard transform by h^*, i.e.

$$h^* = e^{-H^*}. \tag{23}$$

By comparing steps (a), (b), and (c) of the rule above with the method of computing the minimal cut lower bounds for the reliability of the simplified phase configurations, we can conclude immediately that $e^{-H_j^*} \leqslant h_{LBJ}^-$, $j = 1, \ldots, m$. It then follows from (16) and (23) that

$$h^* \leqslant \rho_{PLB-CC} \tag{24}$$

and hence from (18) and (21) that h^* is a lower bound on mission reliability or, equivalently, that the approximate mission hazard transform is conservative.

An algorithm for computing the lower bound h^* follows the first three steps of the algorithm for computing ρ_{PBL-CC}. The next steps are

(e) Compute the component hazards

$$u_{kj} = -\log \rho_{kj}$$

for all (i,j) for which ρ_{kj} has been computed in Step (c).

(f) Compute the approximate mission hazard transform

$$H^* = \sum_{j=1}^{m} \sum_{i=1}^{k(j)} \prod_{C_k \in K_{ji}} u_{kj}.$$

(g) Compute the lower bound

$$h^* = e^{-H^*}.$$

A comparison of this algorithm with the one presented in Section 6 indicates that the computation of the lower bound h^* requires more effort than the computation of the lower bound ρ_{PLB-CC}; we also know from (24) that h^* is less precise than ρ_{PLB-CC}. Thus, it may seem counterproductive to pursue the approximate mission transform any further. However, if one is willing to—or has to, for lack of better information—assume constant component phase failure rates, then the component hazards u_{kj} take on the simple form $u_{kj} = \sum_{i=1}^{j} r_{ki} d_i$, and computations are simplified considerably. In this case, an algorithm for h^* consists of the following steps (expressed in an "intelligible" form):

ALGORITHM 2: for computing h^* in the case of constant component phase failure rates.

(a) Find the minimal cut sets for all phase configurations.

(b) Perform cut cancellation.

(c) Compute the component hazard u_{kj} for each phase j in which component C_k is relevant from

$$u_{kj} = \sum_{i=1}^{j} r_{ki} d_i.$$

(d) Obtain the approximate mission hazard transform from

$$H^* = \sum_{\{\text{all phases}\}} \; \sum_{\substack{\{\text{all min cut sets} \\ \text{in each phase}\}}} \; \prod_{\substack{\{\text{all components in} \\ \text{each min cut set}\}}} u_{kj}$$

(e) compute the lower bound h^* from

$$h^* = e^{-H^*}.$$

When component phase failure rates are assumed constant, the approximate mission hazard transform becomes a polynominal in each of the phase durations. Thus, it is well suited for parametric studies, as is demonstrated in the next example.

EXAMPLE 5. Consider the mission of Example 4. Assume that—all other data being the same as before—the duration of phase 2, d_2, is now uncertain, and that a sensitivity analysis on it is desired.

From the algorithm above, we obtain the following general expression for the approximate mission hazard transform:

$$\begin{aligned} H^* &= r_{31}d_1 r_{51}d_1 \\ &\quad + (r_{11}d_1 + r_{12}d_2) + (r_{31}d_1 + r_{32}d_2(r_{41}d_1 + r_{42}d_2) \\ &\quad + (r_{11}d_1 + r_{12}d_2 + r_{13}d_3)(r_{31}d_1 + r_{32}d_2 + r_{33}d_3) \\ &\quad + (r_{21}d_1 + r_{22}d_2 + r_{23}d_3)(r_{31}d_1 + r_{32}d_2 + r_{33}d_3) \\ &\quad + (r_{21}d_1 + r_{22}d_2 + r_{23}d_3)(r_{61}d_1 + r_{62}d_2 + r_{63}d_3). \end{aligned} \tag{25}$$

H^* as a function of d_2 can be written as

$$H^*(d_2) = a + bd_2 + cd_2^2,$$

with

$$\begin{aligned} a &= d_1 r_{11} \\ &\quad + d_1^2(r_{11}r_{31} + r_{21}r_{31} + r_{21}r_{61} + r_{31}r_{41} + r_{31}r_{51}) \\ &\quad + d_1 d_3(r_{11}r_{33} + r_{13}r_{31} + r_{21}r_{33} + r_{23}r_{31} + r_{21}r_{63} + r_{23}r_{61}) \\ &\quad + d_3^2(r_{13}r_{33} + r_{23}r_{33}, + r_{23}r_{63}) \end{aligned}$$

$$b = r_{12} + d_1(r_{11}r_{32} + r_{12}r_{31} + r_{21}r_{32} + r_{21}r_{62} + r_{22}r_{31} + r_{22}r_{61} + r_{31}r_{42} + r_{32}r_{41}) + d_3(r_{12}r_{33} + r_{13}r_{32} + r_{22}r_{33} + r_{22}r_{63} + r_{23}r_{32} + r_{32}r_{62}),$$

$$c = r_{12}r_{32} + r_{22}r_{32} + r_{22}r_{62} + r_{32}r_{42}.$$

For the data given in Example 4, the numerical values of these coefficients are

$$a = 0.012030$$

$$b = 0.023333 \text{ hours}^{-1}$$

$$c = 0.000258 \text{ hours}^{-2}.$$

For various durations of phase 2 (in hours), the approximate mission hazard transform H^* and the lower bound on mission reliability h^*, both rounded to four decimals, are shown below.

d_2	H^*	h^*
0	0.0120	0.9880
1	0.0356	0.9650
2	0.0597	0.9420
3	0.0844	0.9191
4	0.1095	0.8963
5	0.1351	0.8736
6	0.1613	0.8510
7	0.1880	0.8286
8	0.2152	0.8064
9	0.2429	0.7843
10	0.2712	0.7625

For $d_2 = 2$ hours, ρ_{PLB-CC} and p have been computed in Example 4. Their values are repeated below, together with the value of $h^*(d_2 = 2 \text{ hours})$, to facilitate a comparison.

$$p = 0.9468$$

$$\rho_{PLB-CC} = 0.9438$$

$$h^*(d_2 = 2 \text{ hours}) = 0.9420.$$

In the case of constant component phase failure rates, the approximate hazard transform can also be used to estimate mission reliability when phase durations vary randomly. If $D_1, \ldots, D_m$ are nonnegative random variables denoting the durations of the phases, then the approximate mission hazard transform is $EH^*(D_1, \ldots, D_m)$, where the function H^* is defined as before and E denotes expectation. As an approximation to mission reliability we now use

$$g^* = e^{-EH^*(D_1, \ldots, D_m)}, \tag{26}$$

which is much easier to calculate than the exact value $Ee^{-H^*(D_1, \ldots, D_m)}$. Since e^{-x} is a convex function of x, it follows from Jensen's inequality that $e^{-EH^*(D_1, \ldots, D_m)} \leqslant Ee^{-H^*(D_1, \ldots, D_m)}$, and therefore g^* is a lower bound on mission reliability.

In our last example, we show how this approximation can be used, even without a complete knowledge of the probability distributions of the D_j's.

EXAMPLE 6. Consider again the mission of Example 4, but this time assume that—all other data being the same as before—the durations of phases 2 and 3 are random. The mean durations are known to be $ED_2 = d_2 = 2$ hours and $ED_3 = d_3 = 10$ hours, and the total duration of these two phases together is $D_2 + D_3 = 12$ hours. An estimate for the mission reliability under these circumstances is wanted.

By rearranging the terms of (25) we can express H^* as a function of D_2 and D_3 by

$$H^*(D_2,D_3) = a_1 + a_2D_2 + a_3D_3 + a_4D_2^2 + a_5D_3^2 + a_6D_2D_3,$$

where the coefficients $a_1, \ldots, a_6$ depend only on the known duration of phase 1 and the component phase failure rates. Since $D_2 + D_3 =$ constant, then Var $D_2 =$ VarD_3, and Cov $(D_2,D_3) = -Var\ D_2 = -\text{Var}D_3$. Denoting this common but unknown variance by σ^2, we can write $ED_2^2 = \sigma^2 + d_2^2$, $ED_3^2 = \sigma^2 + d_3^2$, and $ED_2D_3 = d_2d_2 - \sigma^2$, and obtain

$$EH^*(D_2,D_3) = a_1 + a_2d_2 + a_3d_3 + a_4d_2^2 + a_5d_3^2 + a_6d_2d_3 + \sigma^2(a_4 + a_5 - a_6),$$

or, numerically,

$$EH^*(D_2,D_3) = 0.059728 + 0.000071\ \sigma^2/\text{hours}^2.$$

For $\sigma^2 = 0$, i.e. when the durations of phases 2 and 3 take on their expected values with probability one, $EH(D_2,D_3) = 0.0597$ and $g^* = 0.9420$, which agrees with the corresponding results of Example 5. As σ^2 increases, $EH^*(D_2,D_3)$ increases and g^* decreases: since σ^2 cannot be greater than 20 hours^2 under the given conditions, the maximum value of $EH^*(D_2,D_3)$ is 0.0611, and the corresponding minimum value of g^* is 0.9407. We can therefore conclude that the mission reliability is at least 0.94.

ACKNOWLEDGMENTS

This paper summarizes a portion of the author's research at the Naval Postgraduate School. The author acknowledges gratefully the guidance and encouragement of his research advisor, Professor James D. Esary.

REFERENCES

[1] Esary, J.D., and W.J. Hayne, "Properties of an Approximate Hazard Transform," Naval Research Logistics Quarterly *22*, 41-53 (1975).

[2] Esary, J.D., and F. Proschan, "Coherent Structures of Non-Identical Components," Technometrics *5*, 191-209 (1963).

[3] Esary, J.D., F. Proschan, and D.W. Walkup, "Association of Random Variables, with Applications," Annals of Mathematical Statistics *38*, 1466-1474 (1967).

[4] Esary, J.D., and H. Ziehms, "Reliability Analysis of Phased Missions," in *Reliability and Fault Tree Analysis, Theoretical and Applied Aspects of System Reliability and Safety Assessment*, R.E. Barlow, J.B. Fussell and N.D. Singpurwalla, Eds. pps. 213-236, (Society for Industrial and Applied Mathematics, Philadelphia 1975).

[5] Rubin, J.C., "The Reliability of Complex Networks," Proceedings of the AeroSpace Reliability and Maintainability Conference (1964).

[6] Rubinstein, D., "On the Estimation of System Reliability, "Proceedings of the 1961 National Aerospace Electronics Conference, Dayton, Ohio, 262-266 (1961).

[7] Rubinstein, D., "On the Inference of System Reliability, Report I — Mathematical Models," General Electric Radio Guidance Operation, Report R65RG05 (1965).

[8] U.S. Navy Strategic Systems Project Office, *Reliability Evaluation Program Manual* NAVORD OD 29304 Revision A (1974).

[9] Weisberg, S.A., and J.H. Schmidt, "Computer Technique for Estimating System Reliability," *Proceedings of the 1966 Annual Symposium on Reliability,* IEEE 7C26 (1966).

INSPECTION POLICIES FOR DETERIORATING EQUIPMENT CHARACTERIZED BY N QUALITY LEVELS

Z. Kander*

Technion — Israel Institute of Technology
Haifa, Israel

ABSTRACT

Inspection models deal with operating systems whose stochastic failure is detected by observations carried out intermittently. The current communication deals with systems in which $N+1$ levels of quality can be diagnosed. Optimal policies leading to minimal loss are developed, while the system's distribution is represented by an $(N+1)$-state semi-markov process. Based on previous studies of the authors, relative efficiencies of the proposed checking policies are ascertained, by comparison with the loss sustained if discrimination of quality by intermediate levels is disregarded and by viewing the system as one which is either good or failed. Various models are treated where checking, truncated checking, and monitoring policies optimize loss per cycle, per unit of time and per unit of good time. Numerical examples are given.

I. INTRODUCTION

Previous communications [4-6] dealt with operating systems whose stochastic failure was detected by observations carried out intermittently. Three types of costs were introduced: one pertaining to the expense incurred for each check; the second associated with the time elapsing between system failure and its discovery at the subsequent check; and the third relating to replacement of the failed system. The optimal policy is a sequence of checking times $\{t_k,\ k = 1,2,3,...\}$ minimizing the loss per life cycle or, alternatively, per time unit. Models of pure checking, truncated checking, and monitoring additional to checking were treated.

Whereas systems treated before show only two quality characteristics, "good" or "failed", real life abounds with an infinity of cases where quality deteriorates in many levels from "perfectly good" to "totally failed." In effect, for most instances, it can be stipulated that the "good, failed" case is an idealization which comes about by gathering many quality levels into two easily discernible states.

A multilevel quality system in the current study is described by a semi-markov process. The system can move from N, the perfectly good state, only downwards to $N-1$, $N-2$,... until it reaches 0, the failed state. No aging takes place during the stay in any state.

The models of previous studies, as described above, take on a new form when applied to "diagnostic distributions." A time-interval sequence $\{\tau_N,\ \tau_{n-1},\ \ldots,\ \tau_1\}$ describes the policy; i.e.,

*This paper was accepted for publication prior to the death of Dr. Kander in 1973. The manuscript was finalized for publication by Professor Shelemyahu Zacks, Case Western Reserve University.

τ_i, $i=N, N-1, \ldots, 1$ is the periodic checking schedule to be adopted during stay in state i. Consequently in these models, contrary to the previous models (employing good-failed distributions), time from system's operational start is not memorized.

The loss of time's origin constitutes a deficiency of information for the diagnostic distribution, while on the other side the higher number of quality levels is a gain. Numerical examples in later sections indeed demonstrate that relative optimal losses change from model to model to the profit of either distribution.

Various applications of inspection models have been discussed in the literature (see [1], [7], [3], [5]), while multilevel quality systems broaden the field. One example of widespread usefulness of the latter are "redundant structures" [1], like parallel and standby systems which constitute a great part of modern equipment.

II. BASIC ASSUMPTIONS

An inspection model of the present study is well-characterized if three specific sets of assumptions are spelled out:

(i) A statement has to be made about the deterioration mechanism of the unit under study and about the diagnostic power of an observaton carried out on the unit. In this study we shall assume that at any given time the unit is in one of a finite number of feasible states, that a unique (temporal) ordering pertains to these states, that the mechanism of deterioration consists of successive Poisson transitions of the unit from the prevailing state to the consecutive state, and that inspection reveals the prevailing state. The Poisson transition parameters are assumed to be known. A typical example of such situation is that of a standby redundant structure. This is a unit made up of N (not necessarily identical) subunits; the unit is capable of rendering its mission as long as one subunit (at least) has not failed. At any given time, one subunit only is subject to possible random (Poisson) failure. After such failure has taken place, a further subunit is moved up, as it were, and carries out the mission of the unit until it fails, too. This is repeated until failure of the last subunit. Inspection discloses which particular subunit is carrying out the unit mission. Another example would be a structure exhibiting *parallel redundancy*. Here the unit is assumed to be made up of N identical subunits, each of which is subject to Poisson breakdown. The mission of the unit can be carried out as long as one subunit (at least) has not failed. Inspection makes known the number of failed units.

(ii) We shall deal here with three feasible inspection modes:

A) PURE CHECKING. Successive checks are carried out at each check, the state of the system is observed, and a decision about when to check next is made and so on. The failure of the unit terminates one cycle.

B) TRUNCATED CHECKING. Successive checks are carried out; at each check the state of the system is observed. The set of all operative states is subdivided into two subsets, the subset of "early" states and the subset of "late" states. If the check reveals the unit to be in an early state, a decision is made about when to check next. If the system is in a late state, the decision is made to terminate the operation of the unit and a cycle is completed. Another feasible completion of the cycle occurs when it is found in a failed state.

C) CHECKING FOLLOWED BY MONITORING. Here again the set of all operative states is subdivided into two subsets of early and late states, respectively. Successive checks are carried out as before. As long as a check encounters the unit in an early state, a decision is

made about when to check next. If a check finds the unit in a late state, a continuous monitoring procedure is initiated which extends until the failure of the unit and the termination of the cycle without any down time of the unit. Another possibility — possessing nonzero probability — of cycle termination (involving down time) is that the first of two successive checks finds the unit in an early state while the following check reveals failure. It is implicitly assumed that the checking and monitoring procedures bear no influence on the natural life characteristics of the unit.

(iii) Finally, the sources of costs have to be specified and a loss function — whose minimization is the objective of the inspection procedure — has to be set up. We shall assign a cost I to each check that has to be carried out. The cost rate a is assumed to be associated with each unit of down time. Termination of the cycle involves expenditure R if it follows failure. If the cycle is completed through termination of an operative unit, the expenditure incurred is assumed to be equal to S. Monitoring causes two cost components to make their appearance. There is a set-up cost M and, in addition, a cost rate d incurred for each time unit of monitoring. Further costs may show up in actual inspection situations and, in our experience, there is no difficulty in fitting them into our general frame of reference. Under different sets of concrete circumstances we may desire to minimize (at least) three distinct types of loss functions: a) We may be concerned with the events occurring during one single cycle of unit operation. In this case our objective will be to minimize the expected total loss L over the cycle. b) The termination of the cycle may be followed by reactivation of the (identical) unit through replacement or reconstruction. In such cases we assume an infinite horizon to the model and concern ourselves with the minimization of the average loss l per unit of time. c) Again, situations may arise where it is proper to view the stucture of the problem as possessing an infinite horizon, but where it is of interest to minimize expected loss per unit of *good* time (i.e., operation time of the system) rather than cost per unit time at large. The loss function pertaining to this situation will be described by the letter λ.

The notation describing the various combinations of loss functions and inspection modes will be identical with that employed by us in an earlier communication [5]. The letters *c*, *ct*, *cm* refer to pure checking, truncated checking, and checking followed by monitoring, respectively. The letters, L, l, and λ describe what type of overall loss function is to minimized in the context under study. Use of curly brackets is associated with model specification; absence of curly brackets signifies preoccupation with the loss function proper.

III. OPTIMIZATION OVER ONE CYCLE

The unit under investigation passes through N consecutive operational states until it reaches the failed state. For notational convenience we shall assign the subscripts $N, N-1, \ldots, 1$ to the first, second, ... Nth operational states, respectively; the failed state is associated with the subscript 0. A decision procedure is then a set of instructions $[\tau_N, \tau_{N-1}, \ldots, \tau_i, \ldots, \tau_1]$ about when to check again, given that the check presently carried out has revealed the unit to be in state i. Starting from this premise, we obtain the expected future loss during the current cycle K_i as

$$K_i = p_{i,i}(\tau_i)\ K_i + p_{i,i-1}(\tau_i)\ K_{i-1} + \ldots + p_{i,j}(\tau_i)\ K_j + \ldots + p_{i,1}(\tau_i)\ K_1 + p_{i,0}(\tau_i)\ K_0 + J_i(\tau_i)$$

$$K_0 = J_0 \qquad (1 < i \leqslant N,\ 0 \leqslant j < i), \tag{1}$$

where $p_{i,j}(\tau_i)$ is the probability that the system (presently in state i) will be found in state j upon reinspection after time τ_i. The quantity $J_i(\tau_i)$ represents the expected loss which will be incurred by the unit until (and including) the next check; J_0 is the loss that follows a check

revealing the system to be in a failed state. Clearly, this formulation lends itself to the development of an iterative procedure for the purposes of optimization as indeed was already pointed out in Ref. [1]. Relation (1) is transformed in an obvious fashion into

$$K_i = \frac{1}{1-p_{i,i}(\tau_i)} \left[J_i(\tau_i) + \sum_{j=0}^{i-1} p_{i,j}(\tau_i)\ K_j \right] \qquad 0 \leqslant i \leqslant N,\ 0 \leqslant j \leqslant i \tag{2}$$

$$K_o = J_o. \tag{2'}$$

Minimization of the total expected loss function

$$L = K_N \tag{3}$$

is now carried out step by step. We note that

$$K_1(\tau_1) = \frac{1}{1-p_{11}(\tau_1)} \left[J_1(\tau_1) + p_{10}(\tau_1)\ J_o \right], \tag{4}$$

from which expression it is not difficult to derive τ_1^* such that $K_1(\tau_1)$ is brought to its minimal value

$$\min K_1(\tau_1) = K_1^* = K_1(\tau_1^*). \tag{5}$$

Proceeding from here in an analogous manner, we have

$$K_i(\tau_i) = \frac{1}{1-p_{i,i}(\tau_i)} \left[J_i(\tau_i) + \sum_{j=0}^{i-1} p_{i,j}(\tau_i)\ K_j^* \right], \tag{6}$$

and now that value τ_i^* of the feasible τ_i is chosen which minimizes (6), we finally derive

$$\min K_i(\tau_i,\ \tau_{i-1}^*,\ \tau_{i-2}^*,\ \ldots,\ \tau_1^*) = K_i^* = K_i(\tau_i^*,\ \tau_{i-1}^*,\ \ldots,\ \tau_1^*). \tag{7}$$

Overall optimization is attained when the optimal value of the last τ, that is, τ_N^*, has been determined.

To gain better understanding beyond the formalism of the iterative procedure, we shall deal now with various specific cost structures and modes of inspection.

Model {Lc} — Pure Checking

When the system is found in a failed state, some expense (positive or negative) may ensue. Hence it is convenient to set

$$J_o = R \tag{8}$$

where R is typically playing no significant role as long as cycle optimization (rather than optimization in unit time) is required. Introduction of R appears artificial at the present stage — and, indeed, setting $R = 0$ leads to the identical, optimal time sequence — but it is important for further developments.

The expected expenditure until (and including) the next check, given that at present the unit is in state i, is given by

$$J_i(\tau_i) = I + a \int_o^{\tau_i} p_{i,o}(x)\ dx \qquad (0 < i \leqslant N). \tag{9}$$

The right-hand side of (9) is to be interpreted as follows: there is certainly going to follow a check costing I monetary units; furthermore, the integral represents the average time during which the unit is in the failed state 0.

If, now, the quantities J_o, J_1,..., are successively introduced into the loss functions K_1, K_2,..., we finally arrive at the overall expected loss function L equal to

$$Lc = K_N = lE(n) + aE(\Delta t) + R, \tag{10}$$

where $E(n)$ and $E(\Delta t)$ denote the expected number of checks and the expected time in the failed state before detection, respectively. The numerical optimization procedure yields the optimal checking instructions $[\tau_N^*, \tau_{N-1}^*,..., \tau_1^*]$ and the minimal value K_N^* of the overall loss function. It is not difficult to obtain (as byproducts) the values of the expectations $E(n)$ and $E(\Delta t)$ for this case.

It should be noted that in this model (as well as in some others) the inspection policy has no influence on the expectation of active life time of the unit under inspection. The life time of the unit possesses a distribution which is the convolution of N exponential distributions. The expected life time $E(x)$ is simply the sum of the N expectations pertaining to the various states. The expected cycle time $E(t)$ — the average time elapsing from the initiation of the unit up to detection of failure — is (partly) determined by the inspection policy. However, we clearly have the following relation:

$$E(t) = E(x) + E(\Delta t) \tag{11}$$

Addition of a constant to the loss function cannot influence the optimal strategy of inspection. Hence we can modify the loss function (10)

$$\hat{L} = \hat{K}_N = lE(n) + aE(t), \tag{12}$$

and the appropriately modified J-functions are then given by

$$\hat{J}_o = 0 \tag{13}$$

$$\hat{J}_i = l + a\tau_i. \tag{14}$$

This modification turns numerical optimization into a straightforward procedure.

Model {Lct} — Truncated Checking

The present mode of inspection envisages discontinuations of the unit's operation once a certain state has been exceeded. Hence a decision procedure is made up a) of the specification of this state G among the totality of operative states N, $N-1$,..., j ... 2, 1, and b) of a set of instructions $[\tau_N, \tau_{N-1},..., \tau_i,..., \tau_{G+1}, \tau_G]$ about when to check again given that the check presently carried out has revealed the unit to be in state i such that $G \leqslant i \leqslant N$. We recollect that termination brought about the unit's transition into the failed state 0 ensures an expense R, whereas discontinuation of the unit's operation due to its being revealed in one of the states $G-1$, $G-2$,..., 2, 1 is associated with expenditure S.

We are now in a position to set

$$J_o = R \tag{15}$$

$$J_1 = J_2 = ... = J_{G-1} = S \tag{16}$$

$$J_i = I + a \int_o^{\tau_i} p_{i,o}(x)\, dx \qquad (G \leqslant i \leqslant N) \tag{17}$$

and in principle, this set should be used in computational work for all feasible values of G, i.e., $1 \leqslant G \leqslant N$. We note that the case $G = 1$ is identical with pure checking; this is indeed an alternative with which truncated checking should be concerned. The case $G = N$ cannot be considered feasible since the proper interpretation which has to be attached to this is that the unit's operation should never be started. The search for the optimal value of G may be

shortened if we introduce the reasonable conjecture that no local minimum, other than the global minimum, occurs when the minimum total loss is viewed in its dependence on G. The present expression which is analogous to (10) is given by

$$\begin{aligned} Lct = K_N &= IE(n) + aE(\Delta t) + RF + SF^* \\ &= IE(n) + aE(\Delta t) + R + (S-R)F^*, \end{aligned} \tag{18}$$

where F is the probability of terminating in the failed state and $F^*(= 1-F)$ is the probability of discontinuing the unit's operation after finding it in one of the states $G-1$, $G-2,\ldots,$ 2, 1. Equation (18) represents the overall loss function whose minimum minimorum is sought. The various quantities $E(n)$, $E(\Delta t)$, and F depends, of course, on the optimal checking instructions specified by G and $[\tau_N^*, \tau_{N-1}^*,\ldots, \tau_G^*]$. Again the numerical optimization procedure yields not only these instruction and the minimal value K_N^* of the loss function, but also the values of $E(n)$, $E(\Delta t)$ and F for the optimal case.

We note that in this case, unlike model {c}, the expectation of the active life time is a function of the inspection policy pursued. Hence (11) will not be valid under the present circumstances. The expected active life, $E(x')$ say, falls short of $E(x)$ and this quantity too may be obtained within the framework of numerical computation of the optimum policy.

Model {Lcm} — Checking Followed by Monitoring

This mode of inspection pursues the same type of checking as before, i.e., discrete inspections after $\tau_N,\ldots\ \tau_i\ \ldots$ time units until a check reveals that a certain state G has been exceeded. However, if at present the unit is still in one of the operational states $G-1$, $G-2,\ldots,$ 2, 1, a monitoring system is instantaneously put into service involving a set-up cost M and a monitoring cost rate d per unit time; breakdown of the unit is discovered instantaneously if it occurs after the initiation of monitoring. A decision procedure has to specify the state G and give instructions $[\tau_N, \tau_{N-1},\ldots, \tau_i,\ldots, \tau_{G+1}, \tau_G]$ pertaining to rechecking of the unit under inspection, given that the check presently carried out has revealed the unit to be in state i $(G \leqslant i \leqslant N)$.

The functions J_i are now given by

$$J_o = R \tag{19}$$

$$J_i = \begin{cases} M + d\int_o^\infty p_{i,o}(x)\,dx & (0 < i < G) \\ I + a\int_o^{\tau_i} p_{i,o}(x)\,dx & (G \leqslant i \leqslant N) \end{cases} \tag{20}$$

and again all feasible values G, i.e., $1 \leqslant G \leqslant N$, should be introduced and experimented with. The case $G = 1$ is identical with pure checking. We note that the above representation does not cover the alternative of having the monitoring system initiated at the very beginning of the unit's operation. This will be taken care of below.

Again the quantities J_i are introduced into (2) and values of $K_1^*, K_2^*,\ldots, K_N^*$ are derived for prescribed G. We have then a series of nomimal loss functions depending on G:

$$\left\{K_N^*(1),\ K_N^*(2),\ \ldots,\ K_N^*(G),\ \ldots,\ K_N^*(N)\right\}$$

To this series we adjoin

$$K_N^*(N+1) = M + dE(x) + R \tag{21}$$

which represents the case of pure monitoring. The minimum minimorum is selected and the optimal decision rule is defined. The overall loss function in a representation analogous to (10) and (18) is given by

$$Lcm = IE(n) + aE(\Delta t) + R + MF^* + dE \quad \text{(monitored time)} \tag{22}$$

where F^* is the probability of the monitoring system being activated and E (monitored time) is shorthand for the average duration of the monitoring activity. All the quantities appearing on the right-hand side of (22) are obtained as computational byproducts when K_N^* and $[\tau_N^*, \tau_{N-1}^*, \ldots, \tau_G^*]$ are calculated.

Since the montoring policy does not affect the active life time of the unit, the simple relation (11) holds again as in Model $\{c\}$.

IV. OPTIMIZATION WITH INFINITE HORIZON

In many, if not most, applications the unit is replaced or reconstructered after failing, and the inspection process continues toward an infinite horizon. The loss function that one wishes to minimize in such cases is typically not that which was considered in the preceding section, but rather a function representative of this indefinite continuation. Several choices of objective functions present themselves; we shall dwell on two reasonable alternatives, proceed in detail with respect to one of them, and outline how to go about the analysis with respect to the other. These analyses are based on the approach taken in the preceding Section, and make use of the L-minimization technique employed. Optimization over one cycle is then not only an end at which to aim under a given (possibly not widely occurring) set of circumstances. Rather, this type of optimization serves also as an essential link leading to further distinct methods of obtaining optimal inspection procedures.

The proper choice of the objective function should be made on the basis of an analysis of the concrete situation under study. The multiplicity of possible objective functions is not meant to suggest that they should be used on a trial basis, as it were. Rather the salient point is the close connection between the computational techniques and procedures leading to the identification of optimal rules for diverse objective functions. L-optimization is the gateway through which one can proceed with ease to further, and possibly more realistic, modes of optimization.

The two objective functions to be considered here are (a) the average loss per unit of time, and (b) the average loss per unit of good time (active life time). The expected length of the cycle has been denoted by $E(t)$. Hence the average loss per unit time, l, equals

$$l = \frac{L}{E(t)}. \tag{23}$$

The expected active life $E(x')$ either falls short of, or equals, the expected life time $E(x)$ depending on whether the inspection policy (possibly) terminates an active unit or not. At any rate the average loss per unit of good time is expressed by

$$\lambda = \frac{L}{E(x')}. \tag{24}$$

It follows then that λ-optimization is equivalent, in a straightforward fashion, to L-optimization whenever $E(x')$ is not shortened by inspection. Of the models discussed in the present communication, this is the case for pure checking and for checking followed by monitoring.

In the remainder of this section we shall primarily deal with l-optimization. An iteration method, originally employed by Brender [2] and utilized by the present authors in Refs. [4] and [5], is effected in the following manner. Let a function D be defined

(25) $$D = L - l_1 E(t),$$

where l_1 is an arbitrarily selected trial value of the loss per unit time. For given l_1 and for known component costs we may select that inspection procedure which minimizes the function D. Let this minimal value of D, with trial value l, be denoted by

(26) $$D_1(l_1) = L(l_1) - l_1 E(t|l_1)$$

Next we devise a new trial value l_2 by setting

(27) $$l_2 = \frac{L(l_1)}{E(t|l_1)} = \frac{D_1(l_1)}{E(t|l_1)} + l_1.$$

The new value l_2 is now inserted in (25) and the function D (with presently prescribed l_2) is minimized under the new conditions. Repeating this procedure, we have that at any given stage, after minimization, it is possible to write

(28) $$D_j(l_j) = L(l_j) - l_j E(t|l_j)$$

where D_j is the minimal value of D obtained after the jth iteration. If this function has reached the value 0 (or a number sufficiently close to zero), the iteration procedure stops. It may be shown that attaining a minimal D equal to zero is in keeping with the objective of l-optimization. Convergence of this procedure is ensured as demonstrated in Ref. 2. The value l_j which appears in the final iteration is the optimal (minimal) expected loss per unit time; the checking time sequence leading to minimal $D_j = D$ is the appropriate checking time sequence optimizing l. In order to "translate" properly an objective function of the L-type into an objective function of D-type (and effect l-optimization), we modify the J-functions appearing in the preceding section by deducting $l_j \tau_i$ (for each i) whenever either a or d make their appearance.

λ-optimization, whenever it is not trivially equivalent to L-optimization, is typically derivable from L-optimization in a similar fashion.

It is of some interest to recall a general relation connecting optima of L with optima of l. It holds if the inspection policy does not effect a change in the active life time of the unit under surveillance. Consider an l-optimization problem associated with cost a of lost unit time; let the optimal solution be denoted by $l^*(a)$. Next, consider the corresponding L-optimization problem, possessing, though different cost, α per unit of lost time, to wit

(29) $$\alpha = a - l^*(a).$$

The inspection policy minimizing $L(\alpha)$ is identical with the inspection policy minimizing $l(a)$. Furthermore, the two optima are related to each other through the following equivalent formulas:

(30) $$l^*(a) = \frac{L^*(a - l^*(a))}{E(x)}$$

and

(31) $$L^*(\alpha) = E(x)\, l^*\left[\alpha + \frac{L^*(\alpha)}{E(x)}\right].$$

A proof of this may be found in our previous study [4].

A further problem of some interest is the establishment of bounds pertaining to the optimal loss. There are no new problems in the present context and the interested reader is referred to our previous study [5] as well as to some earlier studies [2, 4].

V. AN EXAMPLE

As an illustrative example, consider a unit possessing a standby redundant structure made up of two identical subunits whose life times are exponentially distributed. We start out with both subunits alive; one serves as a spare part, as it were, and the other one is in a state of readiness and is subject to random failure governed by its life-time distribution. This is state 2 in the terminology of previous sections. If now the ready subunit fails, the other one — the spare part — is instantaneously activated, subject as of now to random failure, and the unit as a whole has moved into state 1. Failure of the second subunit is equivalent to breakdown of the unit and transfer to failed state 0. If now the inspection policy is one of pure checking, there is need to specify two times, τ_2 and τ_1. Reinspection is called for τ_2 time units after the present inspection if the unit has been found in state 2; and diagnosis revealing the unit to be in state 1 will lead to rechecking after τ_1 time units. The cost components are I monetary units per check, a monetary units per unit of lost time and R monetary units as an exit fee — for reconstruction, reacquisition, or (possibly) as a salvage cost. We seek optimization with respect to one cycle, i.e., minimization of the function

$$Lc(I,\ a,\ R) = IE(n) + aE(\Delta t) + R.$$

We recall that each subunit possesses life-time density of the negative exponential type

$$f(x) = \xi e^{-\xi x} \quad x \geqslant 0 \quad \xi > 0. \tag{32}$$

The transition probabilities from state to state, i.e., the probabilities of finding the unit in state j upon reinspecting it after τ time units, given that the present state is i ($i = 1, 2$; $j = 0, 1, 2$; $j \leqslant i$), are obtained as

$$\begin{aligned} p_{2,2}(\tau) &= e^{-\xi\tau} \\ p_{2,1}(\tau) &= \xi\tau e^{-\xi\tau} \\ p_{2,0}(\tau) &= 1 - e^{-\xi\tau}(1+\xi\tau) \\ p_{1,1}(\tau) &= e^{-\xi\tau} \\ p_{1,0}(\tau) &= 1 - e^{-\xi\tau}. \end{aligned} \tag{33}$$

These transition probabilities have to be applied, in principle, to (6), and for the determination of J_i in (6), use has to be made of (9). Now it has been stated before than an equivalent procedure is to apply the probabilities to (12) and determine functions $\hat{J}_i$ through utilizing (14) and (13).

The expected active life is, of course, equal to

$$E(x) = 2/\xi \tag{34}$$

and we aim to minimize

$$\begin{aligned} L &= IE(n) + aE(\Delta t) + R = IE(n) + aE(t) - aE(x) + R \\ &= \hat{L} - aE(x) + R. \end{aligned} \tag{35}$$

Clearly, a procedure minimizing $\hat{L}$ minimizes L at the same time. Furthermore, we have

$$\begin{aligned} \hat{J}_0 &= 0 \\ \hat{J}_1 &= I + a\tau_1 \\ \hat{J}_2 &= I + a\tau_2 \end{aligned} \tag{36}$$

We proceed now in the following manner. First we set

(37)
$$\hat{K}_o^* = \hat{K}_o = \hat{J}_o = 0$$
$$\hat{K}_1^* = \min\left\{\frac{\hat{J}_1(\tau_1) + p_{1,0}(\tau_1)\,\hat{K}_o^*}{1 - p_{1,1}(\tau)}\right\}$$
$$= \min\left\{\frac{I + a\tau_1}{1 - e^{-\xi\tau_1}}\right\}$$

(38)
$$= \frac{I + a\tau_1^*}{1 - e^{-\xi\tau_1^*}}.$$

A numerical procedure, such as one carried out on a computer, will yield the optimal values τ_1^* and $\hat{K}_1^*$ if the appropriate computational optimization routine is utilized. However, (38) is of such simple structure that differentiation and setting $dK_1^*/d\tau_1$ equal to 0 leads immediately to the desired aim. We obtain after some manipulation

(39)
$$\frac{I\xi}{a} = e^{\xi\tau_1^*} - (1+\xi\tau_1^*),$$

so that rather elementary mathematical tables may be used for the numerical determination of τ_1^*. If the value of τ_1^* is inserted in (38), it is easily established that

(40)
$$\hat{K}_1^* = I + a\left(\tau_1^* + \frac{1}{\xi}\right).$$

The values of τ_2^* and of $\hat{K}_2^*$ are determined in an analogous manner. We have that

$$\hat{K}_2^* = \min\left\{\frac{\hat{J}_2(\tau_2) + p_{2,1}(\tau_2)\,\hat{K}_1^* + p_{2,0}(\tau_2)\,\hat{K}_o^*}{1 - p_{2,2}(\tau_2)}\right\}$$

$$= \min\left\{\frac{I + a\tau_2 + \xi\tau_2 e^{-\xi\tau_2}\,\hat{K}_1^*}{1 - e^{-\xi\tau_2}}\right\}$$

(41)
$$= \frac{I + a\tau_2^*}{1 - e^{-\xi\tau_2^*}} + \frac{\xi\tau_2^* e^{-\xi\tau_2^*}(I + a\tau_1^*)}{(1-e^{-\xi\tau_2^*})\,(1-e^{-\xi\tau_1^*})}$$

Again, in this particular set of circumstances, a convenient way of deriving the optimal values τ_2^* and $\hat{K}_2^*$ is through the differentiation of $\hat{K}_2$ with respect to τ_2 and setting the derivative equal to 0. Use of (34) and (35) leads to

(42)
$$L^* = \hat{K}_2^* - \frac{2a}{\xi} + R.$$

Since furthermore we have from previous considerations that

(43)
$$\hat{L} = \hat{K}_2 = IE(n) + aE(t),$$

we are able to obtain the expectations $E(n)$ and $E(t)$ through the determination of the multipliers accompanying I and a:

$$E(n) = \frac{1}{1 - p_{2,2}(\tau_2)} \left[1 + \frac{p_{2,1}(\tau_2)}{1 - p_{1,1}(\tau_1)}\right]$$

$$= \frac{1}{1 - e^{-\xi\tau_2}} \left[1 + \frac{\xi\tau_2 e^{-\xi\tau_2}}{1 - e^{-\xi\tau_1}}\right] \tag{44}$$

$$E(t) = \frac{1}{1 - p_{2,2}(\tau_2)} \left[\tau_2 + \frac{\tau_1 p_{2,1}(\tau_2)}{1 - p_{1,1}(\tau_1)}\right]$$

$$= \frac{1}{1 - e^{-\xi\tau_2}} \left[\tau_2 + \frac{\xi\tau_1\tau_2 e^{-\xi\tau_2}}{1 - e^{-\xi\tau_1}}\right] \tag{45}$$

The present development gives only the relatively weak result that (44) and (45) hold under optimal conditions. However, purely probabilistic argumentation shows that they are valid (and, indeed, this is what is stated in (44) and (45), stars having been omitted) for all procedures (τ_1, τ_2).

VI. SOME NUMERICAL RESULTS AND COMPARISONS

The present communication is concerned with the utilization of information acquired when a check is carried out. A measure of the value of the information available in such a fashion would be the difference in losses associated with two distinct optimal inspection schemes: one that utilizes the *diagnosis* of the actual state in which the unit is found on inspection and the other that takes cognizance of part of the information only, to wit, the unit's age.

As an example, consider a standby redundant structure consisting of N identical subunits. The life time of each subunit is exponentially distributed. We have then that overall lifetime is Erlang-distributed with N stages. The optimal diagnostic procedure in this case will count the number of failed subunits at the time an inspection is carried out; the determination of the next checking time will depend on this information. The ordinary (nondiagnostic) procedure, outlined in our communications [4-6], will register only whether or not the unit at large is still active and, if active, will use the present age of the structure — but no more — in order to determine the next checking time.

The general expression for the overall life-time distribution $F(x)$ is obtained as

$$F(x) = p_{N,0}(x). \tag{46}$$

After numerous numerical calculations regarding diagnostic distributions represented by redundant structures of N equal subunits and the respective overall life-time distributions of Erlang type, the following inferences seem justified:

(a) Curtailed inspection, i.e., procedure $\{ct\}$, and monitoring after checking, i.e., procedure $\{cm\}$, are more economic than pure checking policy $\{c\}$.

(b) The $\{c\}$ procedure of pure checking leads to smaller loss for a diagnostic distribution than for the respective overall distribution.

(c) Higher economy for the diagnostic distribution than for the overall simple one seems attainable also in the case of $\{lct\}$ and $\{lcm\}$ modes but not so for $\{Lct\}$ and $\{Lcm\}$ procedures.

Several numerical results for the diagnostic distribution of a unit composed of $N = 2$, $N = 10$ equal subunits, and the respective Erlang-type distribution of degree 2 and 10, are depicted in Table 1 and Fig. 1.

TABLE 1 — *Numerical Results for Some Models and Distributions*
$E(x) = 1$, $l = 1$. D_r = Diagnostic (Standby with r equal subunits).
G_r = Gamma (Erlang type) of r-th degree.

Dist.*	Model	a	S	d	M	L/l	$E(n)$	$E(x^2)$		$E(t)$	$E(\Delta t)$	τ_{10}	τ_9	τ_8	τ_7	τ_6	τ_5	τ_4	τ_3	τ_2	τ_1	G	N	t_N	t_e
D_2	Lc	5				3.461	2.085	1		1.275	.2753									.7509	.3873				
		25				6.844	3.884	1		1.118	.1184									4139	.1818				
	Lct	5	1			2.534	1.426		.7401	0.8556	.1156									.5957		2			
		25	5			6.399	1.810		.6744	0.7276	.0533									4020		2			
		25	25			6.844	≡		Lc																
	Lcm	5		2.5		2.663	1.422	1		1.120	.1201									6075		2			
	*l*c	5				2.613	1.667	1		1.396	.3964									9651	.5381				
	*l*ct	5	1			2.613	≡		*l*c																
	*l*cm	5		2.5		2.309	1.260	1		1.200	.1995									.7893		2			
D_{10}	Lc	5				2.546	1.646	1		1.180	.1800	0.9949	9062	.8130	.7272	.6386	.5262	.4494	.3517	.2558	.1507				
		25				4 690	2.816	1		1.075	.0750	0.7318	6578	.5883	.5070	4376	.3636	.2942	.2202	.1507	.0785				
	Lct	5	1			1.926	1.001		.7264	0.7587	.0318	0.7582										10			
	Lcm	5		2.5		1.807	1.034	1		1.054	.0539	0.8323	.7391	.6504	.5646							7			
		5		2.5	1	2.369	1.089	1		1.109	.1087	0.9519	.8633	.7775	.6888	.5957	.5025					5			
												τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	τ_7	τ_8	τ_9	τ_{10}				
G_2	Lc	5				3.548	2.086	1		1.292	.2924	0.7767	4498	.4630	.4454	4346	.4268	.4203	.4134	4040	.3885		11	5.033	5.033
	Lct	5	1			1.945	1.000		.2865	0.3000	.0135	0.3000											1	0.3000	5.000
	Lcm	5		2.5		2.614	1.000	1		1.082	.0819	0.6000											1	0.6000	5.000
	*l*c	5				2.641	1.639	1		1.425	4249	1.009	.6718	.6282	.6070	.5923	.5770	.5507	.4903				8	5.126	5.126
	*l*ct	5	1			2.641	≡		*l*c																
	*l*cm	5		2.5		2.343	1.000	1		1.163	.1634	0.8000											1	0.8000	5.000

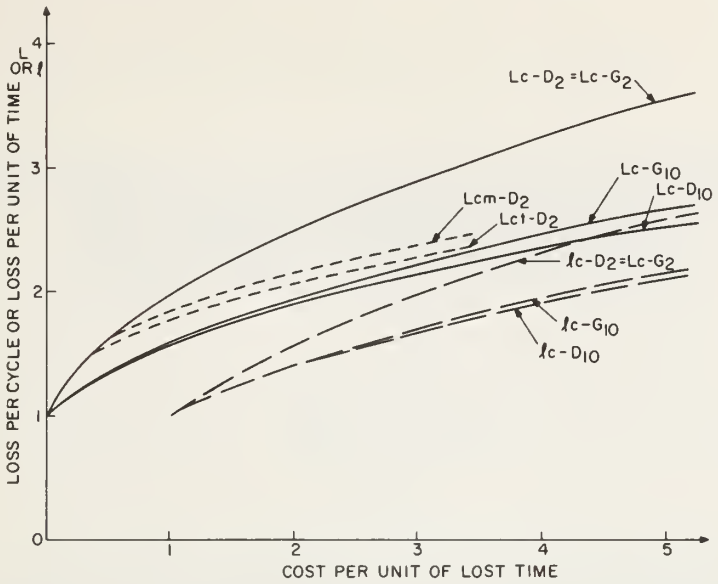


FIGURE 1. Losses L and l for different models. $E(x)=1$, $l=S=1$, $R=M=0$, $d=2.5$. D_i — Diagnostic standby redundant structure of i equal subunits, G_i — Gamma distribution if i-th degree.

REFERENCES

[1] Barlow, R. E., and F. Proschan, *Mathematical Theory of Reliability,* (John Wiley, New York 1965).

[2] Brender, D. M., "A Surveillance Model for Recurrent Events," IBM Watson Research Center Report (1963).

[3] Jorgenson, D. W., J. J. McCall, and R. Radner, "Optimal Maintenance of Stochastically Failing Equipment," The Rand Corporation, R-437-PR (1966).

[4] Kander, Z., and P. Naor, "Optimization of Inspection Policies by Classical Methods," Proceedings of the Third Annual Conference on Operations Research, 1969 (Gordon and Breach, New York, 1970).

[5] Kander, Z., and P. Naor, "Optimization of Inspection Policies by Dynamic Programming," Technion — Israel Institute of Technology, Operations Research, Statistics and Economics, Mimeograph Series No. 61 (Revised Version) (1970).

[6] Luss, H., and Z. Kander, "Inspection Policies When Duration of Checkings is Nonnegligible," Technion Institute of Technology, Operations Research, Statistics and Economics, Mimeograph Series No. 81 (1971).

[7] McCall, J. J., "Maintenance Policies for Stochastically Failing Equipment: A Survey," Management Science *11*, 493-524 (1965).

THE ERROR IN THE NORMAL APPROXIMATION TO THE MULTINOMIAL WITH AN INCREASING NUMBER OF CLASSES

Lionel Weiss*

Cornell University
Ithaca, New York

ABSTRACT

In an earlier paper, it was shown that under certain conditions, if the number of classes in a multinomial distribution increases as the number of trials increases, the probabilities assigned to arbitrary regions by the multinomial distribution are close to the probabilities assigned by the distribution of slightly rounded-off normal random variables. A different method of studying the approximation of the multinomial distribution by a normal distribution is to use the multivariate Berry-Esseen bound. In this paper, these two methods are compared, particularly with respect to the class of multinomial distributions for which the bounds on the error remain useful.

1. EXTENSION OF AN EARLIER RESULT

We briefly review the discussion in Ref. [4], with slightly different notation. For each positive integer n, $\{X_1(n), \cdots, X_{k(n)}(n)\}$ have a joint multinomial distribution, with parameters n, $\{p_1(n), \cdots, p_{k(n)}(n)\}$, where $p_i(n) > 0$, $\sum_{i=1}^{k(n)} p_i(n) = 1$, $\sum_{i=1}^{k(n)} X_i(n) = n$. We assume the following:

(1.1) For some $\Delta > 0$, $\min_{1 \leq i \leq k(n)} [1-p_i(n)] > \Delta$;

(1.2) $\sum_{i=1}^{k(n)} [np_i(n)]^{-1/2}$ approaches zero as n increases;

(1.3) $k(n)[np_{k(n)}(n)]^{-1/2}$ approaches zero as n increases.

(In Ref. [4], it was explicitly assumed that $\min_{1 \leq i \leq k(n)} [np_i(n)]$ approaches infinity as n increases. This fact follows from (1.2).)

Define $Y_i(n)$ as $[np_i(n)]^{-1/2}[X_i(n)-np_i(n)]$, for $i = 1,\ldots,k(n)$, and denote $P[Y_i(n) = y_i;$ $i = 1,\ldots,k(n)-1]$ by $h_n(y_1,\ldots,y_{k(n)-1})$.

Suppose $\{Z_1(n),\ldots, Z_{k(n)-1}(n)\}$ have the following joint normal probability density function:

*Research supported by NSF Grant No. MCS76-06340.

$$\left(\frac{1}{2\pi}\right)^{1/2[k(n)-1]} [p_{k(n)}(n)]^{-1/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{k(n)} z_i^2\right\},$$

where $z_{k(n)}$ is given by the identity $\sum_{i=1}^{k(n)} \sqrt{p_i(n)}\, z_i = 0$. In accordance with this, the random variable $Z_{k(n)}(n)$ is defined by the identity $\sum_{i=1}^{k(n)} \sqrt{p_i(n)}\ Z_i(n) = 0$. We note that $E\{Z_i(n)\} = 0$, Variance $\{Z_i(n)\} = 1 - p_i(n)$, and Covariance $\{Z_i(n), Z_j(n)\} = -\sqrt{p_i(n)p_j(n)}$ for $i,j = 1, \ldots, k(n), i \neq j$.

For $i = 1,\ldots,k(n)-1$, define the random variable $\bar{Z}_i(n)$ as the closest value to $Z_i(n)$ which makes $np_i(n) + \sqrt{np_i(n)}\ \bar{Z}_i(n)$ an integer (positive, negative, or zero). $\bar{Z}_{k(n)}(n)$ is given by the identity

$$\sum_{i=1}^{k(n)} \sqrt{p_i(n)}\,\bar{Z}_i(n) = 0.$$

Define $\bar{\theta}_i(n)$ by

$$\bar{Z}_i(n) = Z_i(n) + \frac{\bar{\theta}_i(n)}{2\sqrt{np_i(n)}}.$$

Then $|\bar{\theta}_i(n)| < 1$ for $i=1, \ldots, k(n)-1$, and $\sum_{i=1}^{k(n)} \bar{\theta}_i(n) = 0$, so $|\bar{\theta}_{k(n)}(n)| \leqq k(n) - 1$. Denote $P[\bar{Z}_i(n) = y_i; i = 1,\ldots, k(n)-1]$ by $g_n(y_1,\ldots, y_{k(n)-1})$.

For any measurable region S_n in $(k(n)-1)$-dimensional space, let $P_{h_n}(S_n)$, $P_{g_n}(S_n)$ denote the probabilities assigned to S_n by h_n, g_n, respectively. In Ref. [4], it was shown that $\lim_{n\to\infty} |P_{h_n}(S_n) - P_{g_n}(S_n)| = 0$ for any sequence $\{S_n\}$. Denote $\max\{k(n)[np_{k(n)}(n)]^{-1/2}, \sum_{i=1}^{k(n)} [np_i(n)]^{-1/2}\}$ by D_n. In this section, we will show that there is a finite value c such that $|P_{h_n}(S_n) - P_{g_n}(S_n)| \leqq c\sqrt{D_n}$.

Denote $\dfrac{h_n(\bar{Z}_1(n),\ldots,\bar{Z}_{k(n)-1}(n))}{g_n(\bar{Z}_1(n),\ldots,\bar{Z}_{k(n)-1}(n))}$ by R_n. If B is any event, let $\bar{B}$ denote its negation. Let Φ denote the standard normal cumulative distribution function. We will use the following elementary inequalities:

(1.4) For any events $B_1, \ldots, B_m$, $P(B_1 \cap \ldots \cap B_m) \geqq 1 - \sum_{i=1}^{m} P(\bar{B}_i)$;

(1.5) For any $x > 0$, $\Phi(x) > 1 - \dfrac{1}{x}\, e^{-x^2/2}$;

(1.6) If Q is a random variable with $P[Q < 0] = 0$, then $P[Q \leqq \sqrt{E(Q)}] \geqq 1 - \sqrt{E(Q)}$.

The inequality (1.5) is a simplified version of a familiar inequality in Ref. [1].

Let A_n denote the event $\{np_i(n) + \sqrt{np_i(n)}\ \bar{Z}_i(n) > \frac{1}{2}\, np_i(n);\ i = 1,\ldots, k(n)\}$. This is the same as the event

$$\left\{ \frac{Z_i(n)}{\sqrt{1-p_i(n)}} > -\frac{1}{2}\sqrt{\frac{np_i(n)}{1-p_i(n)}}\left[1 + \frac{\bar{\theta}_i(n)}{np_i(n)}\right];\ i = 1,\ldots,\ k(n) \right\},$$

and for all sufficiently large n, this event is implied by the event

$$\left\{ \frac{Z_i(n)}{\sqrt{1-p_i(n)}} > -\frac{1}{4}\sqrt{\frac{np_i(n)}{1-p_i(n)}};\ i = 1,\ldots,\ k(n) \right\}.$$

If we use (1.4) and (1.5), the probability of this last event is easily seen to be at least $1 - o(D_n)$, so $P(A_n) \geqq 1 - o(D_n)$.

If A_n occurs, the quantity θ_i appearing in the expression $\Delta(n)$ on p. 145 of Ref. [4] is actually between 0 and 1, since it sets the point at which the third derivative is evaluated in the Taylor's expansion of $\log\left[1 + \frac{\bar{Z}_i(n)}{\sqrt{np_i(n)}}\right]$, and it follows that $\left|\left[1 + \frac{\theta_i \bar{Z}_i(n)}{\sqrt{np_i(n)}}\right]^{-3}\right| \leqq 8$. Examining $\Delta(n)$ on p. 145 of Ref. [4] and $\epsilon_n(\bar{Z}_1,\ldots,\ \bar{Z}_{k(n)-1})$ on p. 146 of Ref. [4], we can easily see that if A_n occurs, then $|\log R_n| \leqq c_1 D_n + c_2 D_n|\ \bar{Z}_{k(n)}(n)\ | + c_3 \sum_{i=1}^{k(n)} [np_i(n)]^{-1/2}$ $\left\{\sum_{j=1}^{4} |\bar{Z}_i(n)|^j\right\} \equiv Q_n$, say, where c_1, c_2, c_3 are fixed finite positive values. From the discussion on 145 of [4], we have $E\{Q_n\} \leqq \bar{c} D_n$ for some finite positive $\bar{c}$. If we use (1.6), it follows that $P[Q_n \leqq \sqrt{\bar{c} D_n}] \geqq 1 - \sqrt{\bar{c}}\ D_n$.

Thus the occurrence of $\{A_n \cap [Q_n \leqslant \sqrt{\bar{c}\ D_n}]\}$ implies that $|\log R_n| \leqslant \sqrt{\bar{c}\ D_n}$. If we use (1.4), it follows that $P[|\log R_n| \leqslant \sqrt{\bar{c}\ D_n}] \geqslant 1 - o(D_n) - \sqrt{\bar{c}\ D_n}$, or $P[e^{-\sqrt{\bar{c}\ D_n}} \leqslant R_n \leqslant e^{\sqrt{\bar{c}\ D_n}}] \geqslant 1 - o(D_n) - \sqrt{\bar{c}\ D_n}$. This last inequality implies that we can find two finite positive constants c_1^*, c_2^*, such that $P[|R_n - 1| \leqq c_1^* \sqrt{D_n}] \geqq 1 - c_2^* \sqrt{D_n}$, and from this, using the argument on pp. 261 and 262 of Ref. [3], we get that $|P_{h_n}(S_n) - P_{g_n}(S_n)| \leqq 2(c_1^* + c_2^*) \sqrt{D_n}$. This completes the demonstration.

2. THE BERRY-ESSEEN BOUND

Let $G_n(S_n)$ denote the probability assigned to a $(k(n)-1)$-dimensional set S_n by the normal cumulative distribution function corresponding to the normal density given above. If S_n is of the form $\{Y_i(n) \leqq y_i;\ i = 1,\ldots,k(n)-1\}$, then Sazonov's result [2] specializes to the inequality

$$|P_{h_n}(S_n) - G_n(S_n)| \leqq$$

$$\frac{C(k(n)-1)}{\sqrt{n}} \sum_{i=1}^{k(n)-1} \left\{ \frac{(1-p_i(n))\left[1 + \frac{P_i(n)}{p_{k(n)}(n)}\right](1 - 2p_i(n) + 2p_i^2(n))}{\sqrt{p_i(n)\ (1-p_i(n))}} \right\}$$

$C(k(n)-1)$ is not given explicitly, but is known to be greater than $\gamma[k(n)-1]^2$ for some positive value γ.

Sazonov gives an analogous result for the case where S_n is any convex set, with a different $C(k(n)-1)$ which is greater than $\bar{\gamma}\,[k(n)-1]^4$ for some positive $\bar{\gamma}$.

3. COMPARISON OF THE TWO BOUNDS

The bound developed in Section 1 is for the approximation of the probability of an arbitrary set by a discretized normal distribution. The bound in Section 2 is for the approximation of the probability of special sets by a normal distribution. We will compare these bounds for large $k(n)$ and sets S_n of the form $\{Y_i(n) \leqq y_i;\ i = 1,\ldots,k(n)-1\}$.

If all $\{p_i(n)\}$ are of the same order of magnitude (that is, if each $p_i(n)$ is approximately $1/k(n)$), then the Berry-Esseen bound of Section 2 is approximately $\frac{2}{\sqrt{n}}\,[C(k(n)-1)]\,\sqrt{k(n)}[k(n)-1]$, which is greater than $\frac{2\gamma}{\sqrt{n}}\,[k(n)-1]^{3+\frac{1}{2}}$. This bound becomes useless unless $(k(n))\,n^{-\frac{1}{7}}$ approaches zero as n increases.

On the other hand, the bound in Section 1 approaches zero as n increases as long as each $p_i(n)$ is approximately $\frac{1}{k(n)}$ and $(k(n))\,n^{-\frac{1}{3}}$ approaches zero. Thus for $k(n)$ in the interval $(n^{\frac{1}{7}},\ n^{\frac{1}{3}})$, we can use the bound of Section 1 but not the bound of Section 2.

Finally, we convert the approximation of $P_{h_n}(S_n)$ by $P_{g_n}(S_n)$ to an approximation of $P_{h_n}(S_n)$ by $G_n(S_n)$, for sets S_n of the form $\{u_i < Y_i(n) \leqq t_i;\ i = 1,\ldots,\ k(n)-1\}$. Let S_n, $S_n(1)$, $S_n(2)$, $S_n(3)$ denote, respectively, the following sets in $(k(n)-1)$-dimensional space:

$$\{u_i < Y_i(n) \leqq t_i;\ i = 1,\ldots,\ k(n)-1\};$$

$$\{u_i < \bar{Z}_i(n) \leqq t_i;\ i = 1,\ldots,\ k(n)-1\};$$

$$\left\{u_i - \frac{1}{2\sqrt{np_i(n)}} < Z_i(n) \leqq t_i + \frac{1}{2\sqrt{np_i(n)}};\ i = 1,\ldots,\ k(n)-1\right\};$$

$$\left\{u_i + \frac{1}{2\sqrt{np_i(n)}} < Z_i(n) \leqq t_i - \frac{1}{2\sqrt{np_i(n)}};\ i = 1,\ldots,\ k(n)-1\right\}.$$

Then we have:

$$(3.1)\quad |\,P_{h_n}(S_n) - P_{g_n}(S_n(1))\,| \leqq c\sqrt{D_n};$$

$$(3.2)\quad G_n(S_n(3)) < P_{g_n}(S_n(1)) < G_n(S_n(2));$$

$$(3.3)\quad G_n(S_n(2)) - G_n(S_n(3)) \leqq \frac{2}{\sqrt{2\pi\Delta}} \sum_{i=1}^{k(n)-1} [np_i(n)]^{-\frac{1}{2}}.$$

The inequality (3.1) is the main result of Section 1. The inequality (3.2) follows directly from the definitions of Section 1. The inequality (3.3) is derived as follows. Since $S_n(3)$ is a subset of $S_n(2)$, $G_n(S_n(2)) - G_n(S_n(3)) = G_n(S_n(2) \cap \overline{S_n(3)})$. But $S_n(2) \cap \overline{S_n(3)}$ is

$$\bigcup_{i=1}^{k(n)-1} \left\{ \left[u_i - \frac{1}{2\sqrt{np_i(n)}} < Z_i(n) \leqq u_i + \frac{1}{2\sqrt{np_i(n)}} \right] \right.$$
$$\left. \cup \left[t_i - \frac{1}{2\sqrt{np_i(n)}} < Z_i(n) \leqq t_i + \frac{1}{2\sqrt{np_i(n)}} \right] \right\}$$

and so

$$G_n(S_n(2) \cap \overline{S_n(3)}) \leqq \sum_{i=1}^{k(n)-1} P\left[u_i - \frac{1}{2\sqrt{np_i(n)}} < Z_i(n) \leqq u_i + \frac{1}{2\sqrt{np_i(n)}} \right]$$
$$+ \sum_{i=1}^{k(n)-1} P[t_i - \frac{1}{2\sqrt{np_i(n)}} < Z_i(n) \leqq t_i + \frac{1}{2\sqrt{np_i(n)}}].$$

But $P[Z_i(n)$ falls in a given interval of length $L_i] = P\left[\frac{Z_i(n)}{\sqrt{1-p_i(n)}}$ falls in the corresponding interval of length $\frac{L_i}{\sqrt{1-p_i(n)}} \right] \leqq \frac{L_i}{\sqrt{2\pi(1-p_i(n))}} \leqq \frac{L_i}{\sqrt{2\pi\Delta}}$. Inequality (3.3) follows immediately.

It is an immediate consequence of the inequalities that $|P_{h_n}(S_n) - P[u_i < Z_i(n) \leqq t_i;\ i = 1, \ldots, k(n)-1]| \leqq \frac{2}{\sqrt{2\pi\Delta}} D_n + c\sqrt{D_n}$.

REFERENCES

[1] Feller, W., *An Introduction to Probability Theory and its Applications, Volume I* (John Wiley and Sons, Inc., New York, 1958).

[2] Sazonov, V. V., "On the Multidimensional Central Limit Theorem," Sankhya, Series A, *30*, 181-204 (1968).

[3] Weiss, L., "The Asymptotic Joint Distribution of an Increasing Number of Sample Quantiles," Annals of the Institute of Statistical Mathematics *21* (2) 257-263 (1969).

[4] Weiss, L., "The Normal Approximation to the Multinomial with an Increasing Number of Classes," Naval Research Logistics Quarterly, *23* (1) 139-149 (1976).

A RECURSIVE ALGORITHM FOR A SUMMED MULTINOMIAL DENSITY FUNCTION

Raymond K. Fink*

Office of Standards Development
U.S. Nuclear Regulatory Commission
Washington, D.C.

Herbert Moskowitz

Krannert School of Management
Purdue University
West Lafayette, Indiana

ABSTRACT

An algorithm for calculating the probabilities of a summed multinomial density function which is recursive with n (the number of trials) is presented. Having application in inspector error models for auditing and quality control problems with Cartesian product structures, the algorithm is discussed in the context of computing optimal economic sampling plans. Computational experience with the algorithm is presented.

INTRODUCTION

In many statistical decision problems, the probabilities of the *adjudged* or reported experimental outcomes, given the *true* experimental outcomes of an experiment, must be determined for the various experiments considered. For example, in inspector error models in auditing and quality control, the computation of $p_n(y|x)$ — where y and x are the reported and the true number of defectives, respectively, in the n items sampled — is essential for determining optimal economic sampling plans [3].†

More generally, this is an example of the following statistical problem. We wish to estimate the posterior distribution of the states $\theta_1,...,\theta_m$ based on the true experimental outcomes $x_1,...,x_n$, but the x's are unobservable. Observed in their place are $y_1,...,y_n$. Thus, the posterior distribution must be estimated using the y's; i.e., we must estimate $p\{\Theta = \theta_i | Y = y_j\}$, where Θ is a random variable whose domain is $\{\theta_1,...,\theta_m\}$ and Y is a random variable whose domain is $\{y_1,...,y_n\}$. To do this, we need to find $p\{Y = y_j | X = x_k\}$, X being a random variable whose domain is $\{x_1,...,x_n\}$, and then apply Bayes' Theorem as follows:‡

*Formerly School of Aeronautics and Astronautics, Purdue University.

†Conclusive evidence exists that inspectors and inspection procedures are fallible, and that these errors are not sufficiently improved with training to be neglected in such model formulations [1,2,3,4]. Errors are introduced into the inspection process when an item is erroneously classified as either bad (Type-I error) or good (Type-II error).

‡It is assumed that the observables (y) are conditionally independent of θ, given x.

$$p\{\Theta = \theta_l | Y = y_j\} = p\{Y = y_j | \Theta = \theta_l\} \cdot p(\Theta = \theta_l)$$

where

$$p\{Y = y_j | \Theta = \theta_l\} = \sum_{k=1}^{n} p\{Y = y_j | X = x_k\} \cdot p\{X = x_k | \Theta = \theta_l\}$$

and

$$p\{Y = y_j = \sum_{l-1}^{m} p\{Y = y_j | \Theta = \theta_l\} \cdot p\{\Theta = \theta_l\}.$$

To illustrate, in quality control θ would represent the lot fraction defective (lot quality state), x the true number of defectives in the sample, and y the number of observed (reported) defectives in the sample (includes those properly and improperly classified). The computation of $p_n(\theta|y)$ would be used in determining an optimal economic sampling plan; i.e., choosing an optimal sample size n^* and acceptance number c^*. The implementation of this plan would involve (a) drawing a single sample of n^* items from a lot of size N, (b) observing the number of defective items y in the sample (perhaps erroneously due to inspection error), and (c) rejecting the lot if more than c^* defectives are observed (i.e., if $y \leqslant c^*$, accept lot, otherwise reject lot).

In discrete inspector error models, the probabilities $p_n(y|x)$ can be typically calculated using a summed multinomial density function. In this paper, a highly efficient recursive algorithm for the computation of these probabilities is presented. The following notation will be used to develop the algorithm:

n	=	number of trials (sample size)
x	=	true number of failures in n trials
y	=	reported number of failures in n trials
$p_n\{y,x\}$	=	joint probability of reported failures and true failures in n trials (an abbreviated notation for $p_n\{Y = y_j, X = x_k\}$)
$p_n\{y\|x\}$	=	conditional probability of reported failures given the true failures in n trials (an abbreviated notation for $p_n\{Y = y_j \| X = x_k\}$).

SUMMED MULTINOMIAL FORMULATION

If we can view an experiment as consisting of n independent trials, each trial having four mutually exclusive and exhaustive outcomes (1) report a success when a true success occurs; (2) report a failure when a true success occurs; (3) report a success when a true failure occurs; (4) report a failure when a true failure occurs) whose probabilities remain constant for each trial, each element of the matrix $[p_n\{y|x\}]$ can be computed using the following summed multinomial density function (see Appendix A for derivation),

$$(1) \qquad p_n\{j|k\} = \left(\sum_{(n_1,n_2,n_3,n_4)\in N} \frac{n!(p_1\{0|0\})^{n_1}(p_1\{1|0\}^{n_2}(p_1\{0|1\})^{n_3}(p_1\{1|1\})^{n_4}}{n_1!n_2!n_3!n_4!} \right) \div \left(\frac{n}{k} \right) \forall_{j,k}$$

where N is the set of solutions $\{(n_1, n_2, n_3, n_4)\}$, $(\leqslant n + 1)$, to the integer linear system $j = n_2 + n_4$, $k = n_3 + n_4$, $n_i \geqslant 0$, integer, and $\Sigma_{i=1}^{4} n_i = n$. $p_1\{0|0\}$ is the probability that the reported outcome is successful given that the true outcome is successful; $p_1\{1|0\}$ is the proba-

bility that the reported outcome is a failure given that the true outcome is successful (Type-I error); and so on.† n_1, n_2, n_3, and n_4 are the number of times outcomes (1), (2), (3), and (4) occur, respectively, in n trials.

However, the computational complexity of this density function and the fact that most algorithms for determining optimal quality control inspection plans examine successively increasing sample sizes, highlight the value of a recursive formulation of $p_n\{y|x\}$ [5]. In this paper we develop from the basic transition equation a recursive algorithm for computing $p_n\{y|x\}$, for quality control and auditing applications, which circumvents the use of the more computationally laborious summed multinomial density function.

RECURSIVE ALGORITHM

Development

If the experiment is composed of n independent trials, then the probabilities of outcomes for n trials can be determined from the probabilities of outcomes for $(n-1)$ trials and the probabilities for the nth trial. We will assume that the probabilities for each trial are identically distributed, and we can therefore write the step-transition equation:

$$(2)\qquad p_n\{j,\ k\} = p_{n-1}\{j,\ k\}p_1\{0,\ 0\} + p_{n-1}\{j,\ k-1\}p_1\{0,\ 1\} + p_{n-1}\{j-1,k\}p_1\{1,\ 0\} + p_{n-1}\{j-1,\ k-1\}p_1\{1,\ 1\},\quad 0 \leqslant j \leqslant n,\ 0 \leqslant k \leqslant n.$$

Replacing the joint probabilities by the appropriate conditional and marginal probabilities and dividing both sides by $p_n\{x = k\}$ yields:

$$(3)\quad p_n\{j,\ k\} = \frac{n-k}{n}\left[p_{n-1}\{j|k\}p_1\{0|0\} + p_{n-1}\{j-1|k\}p_1\{1|0\}\right] + \frac{k}{n}\left[p_{n-1}\{j|k-1\}p_1\{0|1\} + p_{n-1}\{j-1|k-1\}p_1\{1|1\}\right];\quad 0 \leqslant j \leqslant n,\ 0 \leqslant k \leqslant n.$$

For the case $k = 0$, this becomes

$$(3a)\qquad p_n\{j|k\} = p_{n-1}\{j|k\}p_1\{0|0\} + p_{n-1}\{j-1|k\}p_1\{1|0\},$$

and for $k = n$,

$$(3b)\qquad p_n\{j|k\} = p_{n-1}\{j|k-1\}p_1\{0|1\} + p_{n-1}\{j-1|k-1\}p_1\{1|1\}.$$

For the case $0 < k < n$, we can show inductively that both (3a) and (3b) are true (see Appendix B). Thus (3a) can be used for $0 \leqslant k < n$, and (3b) can be used for $0 < k \leqslant n$. This approach forms the basis for the recursive algorithm presented below, which is more efficient than using (3) because less than half as many computations are required.

Matrix Formulation of the Recursive Algorithm

Given the arrays of probabilities $[p_1\{y|x\}]$ and $[p_{n-1}\{y|x\}]$, we wish to find $[p_n\{y|x\}]$. Let

R = an $(n+1)$ x n array constructed with rows from $[p_{n-1}\{y|x\}]$
Q = an $(n+1)$ x 2 array constructed with rows from $[p_1\{y|x\}]$.

†The values for $p_1\{y|x\}$ would be typically determined empirically using past data on inspector performance.

STEP I: CONSTRUCT THE Q ARRAY. The $(n+1)$ x 2 Q array is constructed from $[p_1\{y|x\}]$. The first m rows of Q will be identical to row 1 of $[p_1\{y|x\}]$, and the remaining $(n+1-m)$ rows of Q will be identical to row 2 of $[p_1\{y|x\}]$. The value of m is determined in one of two ways, depending on whether n is even or odd: $m = (n/2)+1$ for even n; $m = (n+1)/2$ for odd n. Thus, $q_{i1} = p_1\{0|0\}$ and $q_{i2} = p_1\{1|0\}$ for $i = 1, 2, \ldots, m$, and $q_{i1} = p_1\{0|1\}$ and $q_{i2} = p_1\{1|1\}$ for $i = m+1, m+2, \ldots n+1$.

STEP II: CONSTRUCT THE R ARRAY. The $(n+1)$ x n array R is constructed by taking the mth row (m as calculated in Step I, above) of $[p_{n-1}\{y|x\}]$ and using it as the mth and $(m-1)$th rows of R. The first $(m-1)$ rows of R are identical to the first $(m-1)$ rows of $[p_{n-1}\{y|x\}]$; the $(m+1)$th row of $[p_{n-1}\{y|x\}]$ becomes the $(m+2)$th row of R, and so on. Thus, for $i \leqslant m$, $r_{ij} = p_{n-1}\{j-1|i-1\}$ for $j =, 1,2,3,\ldots,n$ and for $i > m$, $r_{ij} = p_{n-1}\{j-1|i-2\}$ for $j = 1,2,3,\ldots,n$.

STEP III: CALCULATE $[p_n\{y|x\}]$. $[p_n\{y|x\}]$ is then calculated a row at a time as follows:

$$[p_n\{y|x\}]_\alpha = Q_\alpha \cdot \begin{bmatrix} R_\alpha & 0 \\ 0 & R_\alpha \end{bmatrix} \text{ for } \alpha = 1,2,3,\ldots,n,\ n+1, \tag{4}$$

where the subscript α indicates the α-row of the array. If we use the elements of the arrays in place of the row vectors, (4) can also be written as

$$[p_n\{0|\alpha-1\}p_n\{1|\alpha-1\} \cdots p_n\{j|\alpha-1\} \cdots p_n\{n|\alpha-1\}] \tag{4a}$$
$$= [q_{\alpha1}\ q_{\alpha2}] \cdot \begin{bmatrix} r_{\alpha1} & r_{\alpha2} \cdots r_{\alpha j} \cdots r_{\alpha n} & 0 \\ 0 & r_{\alpha1} \cdots r_{\alpha(j-1)} \cdots r_{\alpha(n-1)} & r_{\alpha n} \end{bmatrix}.$$

It can be easily seen that this matrix formulation embodies the recursive equations (3a) and (3b). We can continue recursively in this fashion to find $[p_{n+1}\{y|x\}]$, $[p_{n+2}\{y|x\}]$, ...

EXAMPLE: Consider the problem where

$$[p_1\{y|x\}] = \begin{bmatrix} .9 & .1 \\ .2 & .8 \end{bmatrix}.$$

We wish to find $[p_n\{y|x\}]$ for $n = 2,3,\ldots$. For $n = 2$:

I. $m = (n/2)+1 = 2 \rightarrow Q = \begin{bmatrix} .9 & .1 \\ .9 & .1 \\ .2 & .8 \end{bmatrix}$

II. $m = (n/2)+1 = 2 \rightarrow R = \begin{bmatrix} .9 & .1 \\ .2 & .8 \\ .2 & .8 \end{bmatrix}$

III. $[p_2\{y|x\}]_1 = [.9 \quad .1] \begin{bmatrix} .9 & .1 & 0 \\ .0 & .9 & .1 \end{bmatrix} = [.81 \quad .18 \quad .01]$

$[p_2\{y|x\}]_2 = [.9 \quad .1] \begin{bmatrix} .2 & .8 & 0 \\ 0 & .2 & .8 \end{bmatrix} = [.18 \quad .74 \quad .08]$

$[p_2\{y|x\}]_3 = [.2 \quad .8] \begin{bmatrix} .2 & .8 & 0 \\ 0 & .2 & .8 \end{bmatrix} = [.04 \quad .32 \quad .64].$

Therefore,

$$[p_2\{y|x\}] = \begin{bmatrix} .81 & .18 & .01 \\ .18 & .74 & .08 \\ .04 & .32 & .64 \end{bmatrix}.$$

For $n = 3$:

$$\text{I.} \quad m = (n+1)/2 \rightarrow Q = \begin{bmatrix} .9 & .1 \\ .9 & .1 \\ .2 & .8 \\ .2 & .8 \end{bmatrix}$$

$$\text{II.} \quad m = (n+1)/2 \rightarrow R = \begin{bmatrix} .81 & .18 & .01 \\ .18 & .74 & .08 \\ .18 & .74 & .08 \\ .04 & .32 & .64 \end{bmatrix}$$

If we use (4), in Step III, $[p_n\{y|x\}]$ thus becomes

$$[p_3\{y|x\}] = \begin{bmatrix} .729 & .243 & .027 & .001 \\ .162 & \boxed{.684} & .146 & .008 \\ .036 & .292 & .608 & .064 \\ .008 & .096 & .384 & .512 \end{bmatrix}.$$

We can continue recursively in this fashion for $n = 4,5,\ldots$. Note that the rows of $[p_n\{y|x\}]$ sum to one as they should, since a row represents the probabilities of various reported experimental outcomes for a given true experimental outcome. Identical results for $[p_n\{y|x\}]$ are obtained by calculating the probabilities using the summed multinomial density function. For example, for $p_3\{1|1\}$ we have $j = n_2 + n_4 = 1$, $k = n_3 + n_4 = 1$, and $n = n_1 + n_2 + n_3 + n_4 = 3$, yielding the set $N = (1,1,1,0)$, $(2,0,0,1)$ which gives

$$p_3\{1|1\} = \frac{3!}{\begin{pmatrix} 3 \\ 1 \end{pmatrix}} \left(\frac{(.9)^1(.1)^1(.2)^1(.8)^0}{1!\,1!\,1!\,0!} + \frac{(.9)^2(.1)^0(.2)^0(.8)^1}{2!\,0!\,0!\,1!} \right)$$

$$= .684.$$

Remarks

In its present form, the recursive algorithm assumes constant, identically distributed trial probabilities. In the context of inspector error models, this implies constant inspector error independent of sample size. However, more sophisticated models of inspector error can be easily accommodated by the algorithm.

For example, if inspector error remains constant in each trial (i.e., for each item inspected in the sample) for a given sample size n, but changes with increasing n, we can reflect this exogeneously adjusting $[p_1\{y|x\}]$ for each sample size n, and setting it to $[p_1\{y|x\}]_{(n)}$ (the inspector error probabilities for a single sample out of a sample size n). For each n we must then use the recursive algorithm to regenerate $[p_1\{y|x\}]_{(n)}$, $[p_2\{y|x\}]_{(n)}$, ... , $[p_n\{y|x\}]_{(n)}$; while this slows things down considerably, it is still much faster than the summed multinomial approach (see Computational Experience below).

A computationally simpler model (which is more realistic for inspection error) is one in which the trials are not identically distributed, but change with sample size such that $[p_1\{y|x\}]_{(k)}$, the error probabilities at the kth trial, are constant for any sample size n, but $[p_1\{y|x\}]_{(k)}$ and $[p_1\{y|x\}]_{(k+1)}$ need not be identical. This may reflect the effects of fatigue on the inspector's performance; in the transition equation (2), this implies variable step-transition probabilities. For this case, we can simply adjust $[p_1\{y|x\}]_{(n)}$ as necessary, and calculate $[p_{n+1}\{y|x\}]$ from $[p_n\{y|x\}]$ with virtually no increase in computation time.

COMPUTATIONAL EXPERIENCE†

Using FORTRAN IV and a CDC 6500 computer, we have generated the arrays for sample sizes from $n = 1$ to 200. With the algorithm, these computations require about 60 s of central processor (CPU) time to generate all matrices $[p_n\{y|x\}]$ from $n = 1$ to 150. A summed multinomial solution requires 348s to generate only the matrix $[p_{150}\{y|x\}]$ (Fig. 1). It should be noted that the coding of the algorithm can eliminate the Q and R arrays, and can use temporary row storage vectors ("temp") to achieve a substantial reduction in central memory requirements on the order of n^2 [6].

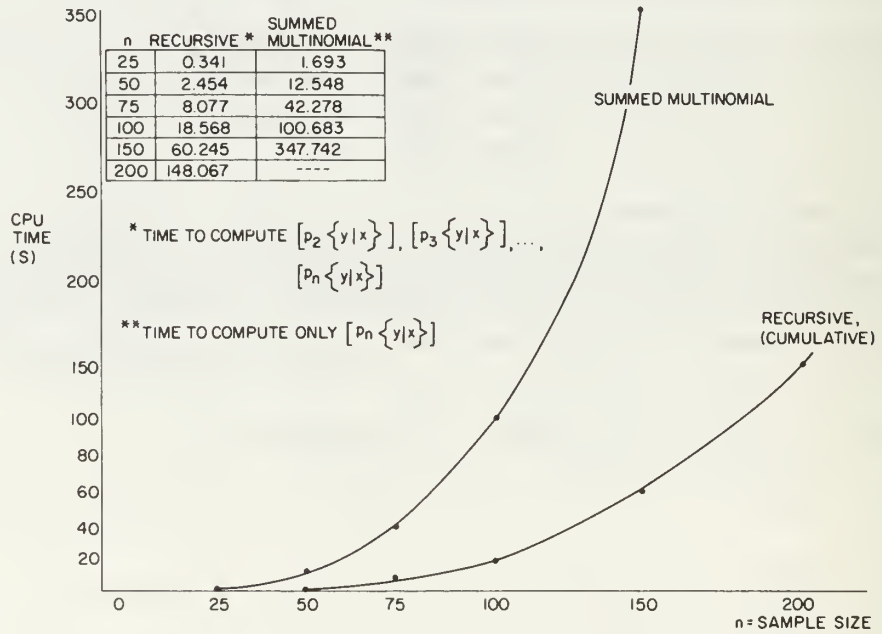


FIGURE 1. — Computational experience; recursive and summed formulations.

CONCLUSION

The algorithm presented describes an efficient method for calculating the probabilities of a summed multinomial density function. The algorithm is particularly advantageous when an array of such probabilities must be computed for a number of different sample sizes as would be the case when attempting to determine an optimal sampling plan for quality control. The

†A computer program abstract and program listing of the recursion algorithm is available on request. The algorithm is used as a subroutine with other computational routines for determining optimal sampling plans in quality control and auditing. Abstracts and listings of these programs are also available on request.

algorithm can also be used to economically develop appropriate tables for a summed multinomial density function, which could be used for constant inspector error models. If inspector error rates are not constant but vary with the numbered item sampled, inspector error tables could also be developed, if a recursion existed between $p_1(j|k)_n$ and $p_1(j|k)_{n+1}$, $j = 0, 1$; $k = 0,1$.

For sample sizes in the range of practical interest, the core memory requirements of the algorithm are not prohibitively large. While the effects of cumulative truncation errors in the algorithm have not been studied in detail, the solution quality has been generally acceptable in applications to date [7]. Moreover, one can resort to extended precision arithmetric to minimize such cumulative truncation errors, however, at the expense of increased core memory usage. The computational advantage provided by the recursive algorithm has enabled the solution of problems which were previously computationally impractical [7].

REFERENCES

[1]. Biegel, J. E., "Inspector Errors and Sampling Plans," AIEE Transaction *6*, 197, 284-287.

[2]. Carroll, H. M., "Estimating Errors in the Inspection of Complex Products," AIIE Transaction, *1*, 229-235 (1969).

[3]. Collins, R. D., "Statistically and Economically Based Attribute Acceptance Sampling Models with Inspection Errors," Ph.D. Dissertation in Industrial Engineering and Operations Research, Virginia Polytechnic Institute and State University (May 1974).

[4]. Collins, R. D., K. E. Case, and G. K. Bennett, "The Effects of Inspector Error on Single Sampling Inspection Plans," International Journal of Production Research, *11*, 289-298 (1973).

[5]. Moskowitz, H. and W. L. Berry, "A Bayesian Algorithm for Determining Optimal Single Sample Acceptance Plans for Product Attributes," Management Science, *22*, (11), 1238-1250 (July 1976).

[6]. Moskowitz, H. and R. K. Fink, "A Program of a Recursive Algorithm for a Summed Multinomial Density Function," Behavior Research Methods & Instrumentation, *7*, (6), 574 (1975).

[7]. Moskowitz, H. and R. K. Fink, "A Bayesian Algorithm Incorporating Inspector Errors for Quality Control and Auditing," in *Algorithmic Methods in Probability*, TIMS Studies in Management Sciences, M. F. Neuts, Ed., Volume 7, pp. 79-104, (North-Holland Publishers 1977).

APPENDIX A

DERIVATION OF THE SUMMED MULTINOMIAL DENSITY FUNCTION FOR DETERMINING $p_n\{y|x\}$

Let $x_1, x_2, \ldots, x_n$ be the possible *actual* or *true* outcomes (data) of an experiment e_t (e.g., sample of size n). Let $y_1, y_2, \ldots, y_n$ be the *adjudged* or *observed* (e.g., inspector or auditor reported) outcomes of an experiment e_t. We wish to find $p_n\{y = j|x = k\}$, for $j, k = 0, \ldots, n$. This can be derived using a summed multinomial density function if: (1) the experiment e_t consists of n identical trials; (2) each trial is a Bernoulli process to which the values 0 and 1 can be assigned. To show this, Let $W_t \epsilon \{0, 1\}$ be the *actual* outcomes of trial t, with $p\{W_t = 0\} = p_1\{0\}$, $p\{W_t = 1\} = p_1\{1\}$, $p_1\{0\} + p_1\{1\} = 1$ $\forall t$ (e.g., let 0 = a nondefective item, 1 = a defective item). Thus, the trials are independent and identically distributed. Let $x = \Sigma_{t=1}^{n} W_t$. Similarly, define $V_t \epsilon \{0, 1\}$ and $y = \Sigma_{t=1}^{n} V_t$ for the *adjudged* outcomes. Let $p\{V_t = 0|W_t = 0\} = p_1(0|0)$, $p\{V_t = 1|W_t = 0\} = p_1(1|0)$, $p\{V_t = 0|W_t = 1\} = p_1(0|1)$,

$p\{V_t = 1 | W_t = 1\} = p_1(1|1)$, be known functions that are fixed $\forall t$. We wish to find $p_n\{y = j | x = k\}$. We know that the four outcomes (events), $\{V_t = 0, W_t = 0\}$, $\{V_t = 1, W_t = 0\}$, $\{V_t = 0, W_t = 1\}$, and $\{V_t = 1, W_t = 1\}$ are mutually exclusive and collectively exhaustive at each trial t, with probabilities $p_1\{0, 0\}$, ..., $p_1\{1, 1\}$, respectively. Let the above four outcomes occur n_1, n_2, n_3, and n_4 times, respectively, in experiment e_t. Then we have: $j = n_2 + n_4$, $k = n_3 + n_4$, $n = \Sigma_{t=1}^{n} n_t$, $n_1, n_2, n_3, n_4 \geqslant 0$, integer, where j, k, and n are known.

The above system has a finite number ($\leqslant n + 1$) of feasible solutions. Let $N = \{(n_1, n_2, n_3, n_4) |$ above linear system satisfied$\}$. Then

$$p_n\{y = j | x = k\} = \frac{p_n\{y = j, x = k\}}{p_n\{x = k\}} \tag{A.1}$$

where

$$p_n\{x = k\} = \binom{n}{k} \cdot [p_n\{0, 1\} + p_n\{1, 1\}]^k \cdot [p_n\{0, 0\} + p_n\{1, 0\}]^{n-k}$$

or

$$p_n\{x = k\} = \binom{n}{k} p_n\{1\}^k \cdot p_n\{0\}^{n-k} \tag{A.2}$$

The joint probability distribution $p_n\{y = j, x = k\}$ can be represented by the summed multinomial density function

$$p_n\{y = j, x = k\} = \Sigma_{(n_1, n_2, n_3, n_4) \in N} \frac{n!(p_1\{0, 0\})^{n_1}(p_1\{1, 0\})^{n_2}(p_1\{0, 1\})^{n_3}(p_1\{1, 1\})^{n_4}}{n_1!\, n_2!\, n_3!\, n_4!}$$

Letting $p_1\{0, 0\} = p_1\{0|0\} \cdot p_1\{0\}$, etc., and substituting and factoring yields

$$p_n\{y = j, x = k\} = \Sigma_{(n_1, n_2, n_3, n_4) \in N} \frac{n!(p_1\{0|0\})^{n_1}(p_1\{1|0\})^{n_2}(p_1\{0|1\})^{n_3}(p_1\{1|1\})^{n_4}}{n_1!\, n_2!\, n_3!\, n_4!} \cdot p_1\{0\}^{n_1+n_2} p_1\{1\}^{n_3+n_4}. \tag{A.3}$$

But $n_3 + n_4 = k$ and $n_1 + n_2 = n - k$. Substituting (A.2) and (A.3) into (A.1) then gives

$$p_n\{y = j | x = k\} = \Sigma_{(n_1, n_2, n_3, n_4) \in N} \frac{n!(p_1\{0|0\})^{n_1}(p_1\{1|0\})^{n_2}(p_1\{0|1\})^{n_3}(p_1\{1|1\})^{n_4}}{n_1!\, n_2!\, n_3!\, n_4!} \div \binom{n}{k} Vj, k \tag{A.4}$$

which is simply a summed multinomial density function, where $j = n_2 + n_4$, $k = n_3 + n_4$, $n = \Sigma_{t=1}^{4} n_t$, $n_t \geqslant 0$, integer, and the matrix $[p_1\{y|x\}]$ is Markov. Note that $(n_1, n_2, n_3, n_4) \in N$ can be obtained by writing n_t in terms of n_4, i.e., $(n-k-j + n_4, j-n_4, k-n_4, n_4) \in N$, all components being nonnegative. Hence,

$$N = \cup_{n_4=0}^{n} \{n-k-j + n_4, j-n_4, k-n_4, n_4) | n-k-j + n_4 \geqslant 0, j-n_4 \geqslant 0, k-n_4 \geqslant 0\}. \tag{A.5}$$

APPENDIX B

THEOREM: For $0 < k < n$,

$$p_n\{j|k-1\} \cdot p_1\{0|1\} + p_n\{j-1|k-1\} \cdot p_1\{1|1\} = p_n\{j|k\} \cdot p_1\{0|0\} + p_n\{j-1|k\} \cdot p_1\{1|0\} \tag{B.1}$$

(and we assume that the above equation holds for $n-1$).

PROOF: Expanding the left hand side of (B.1) with the recursion equation (3) we get

$$\begin{aligned} &p_n\{j|k-1\} \cdot p_1\{0|1\} + p_n\{j-1|k-1\} \cdot p_1\{1|1\} = \\ &p_1\{0|1\} \cdot \frac{n-(k-1)}{n} \cdot [p_{n-1}\{j|k-1\} \cdot p_1\{0|0\} + p_{n-1}\{j-1|k-1\}p_n\{1|0\}] \\ &+ p_1\{0|1\} \cdot \frac{k-1}{n} \cdot [p_{n-1}\{j|k-2\} \cdot p_1\{0|1\} + p_{n-1}\{j-1|k-2\}p_1\{1|1\}] \\ &+ p_1\{1|1\} \cdot \frac{n-(k-1)}{n} \cdot [p_{n-1}\{j-1|k-1\}p_1\{0|0\} + p_{n-1}\{j-2|k-1\} \cdot p_1\{1|0\}] \\ &+ p_1\{1|1\} \cdot \frac{k-1}{n} [p_{n-1}\{j-1|k-2\}p_1\{0|1\} + p_{n-1}\{j-2|k-2\}p_1\{1|1\}]. \end{aligned} \tag{B.2}$$

By induction we assume that the first bracketed term on the right-hand side of (B.2), $[\,]_1$, is equal to the second bracketed term $[\,]_2$, and the third bracketed term, $[\,]_3$, is equal to the fourth, $[\,]_4$. Substituting $[\,]_1$ for $[\,]_2$, and $[\,]_3$ for $[\,]_4$, in (B.1) yields

$$\begin{aligned} &P_n\{j|k-1\} \cdot p_1\{0|1\} + p_n\{j-1|k-1\} \cdot p_1\{1|1\} = \\ &p_1\{0|1\} \cdot [p_{n-1}\{j|k-1\} \cdot p_1\{0|0\} + p_{n-1}\{j-1|k-1\}p_1\{1|0\}]. \\ &+ p_1\{1|1\} \cdot [p_{n-1}\{j-1|k-1\} \cdot p_1\{0|0\} + p_{n-1}\{k-2|k-1\}p_1\{1|0\}]. \end{aligned} \tag{B.3}$$

The identical result in (B.3) would also be obtained by expanding the right-hand side of (B.1) using the recursion equation (3). Hence (B.1) is true.

A PROBABILITY MODEL FOR INITIAL CRACK SIZE AND FATIGUE LIFE OF GUN BARRELS*

F. Proschan and J. Sethuraman

Department of Statistics
The Florida State University
Tallahassee, Florida

ABSTRACT

With constant firing, metal fatigue produces cracks in a gun barrel. The useful life of the barrel comes to an end when a crack develops to a critical size. The theory of Fracture Mechanics suggests a formula for crack size growth rate. This formula can be used to determine the life of a barrel, depending on the initial and critical crack sizes and other factors. The initial crack size turns out to be a dominant factor. Unfortunately, accurate measurements are not generally available on the initial crack size. In this paper, we propose a simple probability model for the initial crack size and this, in turn, leads to a probability distribution of the life of the barrel. This last distribution is the well-known exponential distribution with a location shift. The simplicity of this final result is one of the factors that make the model appealing.

1. INTRODUCTION

Suppose that there is a crack of size b in a gun barrel after N rounds have been fired. The rate at which the crack size will grow, db/dN, is a central topic of study in Fracture Mechanics. The following formula, taken from Davidson, Throop and Reiner[1], represents a reasonably simplified version of many more complicated ones available.

$$\frac{db}{dN} = \frac{C(\Delta K)^m}{E\, S_Y\, K_{Ic}}, \tag{1.1}$$

where

ΔK = increase in stress intensity

E = elastic modulus

S_Y = tensile yield strength

K_{Ic} = fracture toughness

m = a constant between 2 and 4

C = a material constant.

*Research supported by the Army, Navy, and Air Force under the U.S. Army Research Office-Durham grant No. DAAG29-76-G-0238. Reproduction in whole or in part is permitted for any purpose of the U.S. government.

Throop[4] gives a simple expression for ΔK, which depends on the size b and the shape of the crack, as follows:

$$\Delta K = \alpha S \sqrt{\pi b}, \tag{1.2}$$

where

S = bore stress = $p\,\dfrac{\omega^2+1}{\omega^2-1}$,

p = pressure

ω = ratio of outer diameter inner diameter

$$\alpha = \begin{cases} 1.5 \text{ if the crack is a frontal notch} \\ 1.0 \text{ if the crack is semielliptical} \\ 0.5 \text{ if the crack is semicircular.} \end{cases}$$

Thus

$$\frac{db}{dN} = Db^{m/2}, \tag{1.3}$$

where

$$D = [C\alpha^m S^m \pi^{m/2}]/[ESyK_{Ic}]. \tag{1.4}$$

Suppose that the barrel is fired a few rounds to provide a *heat check pattern.* The size of the largest crack produced at this stage is called the *initial crack size* and denoted by b_o. The number of rounds, L, required for this crack to grow to the critical crack size b_c is the *life* of the barrel. From (1.3), it is easy to obtain an expression for L:

$$L = \begin{cases} \dfrac{2}{(m-2)\,D}\,[b_o^{-(m-2)/2} - b_c^{-(m-2)/2}] \text{ if } m > 2. \\ \dfrac{1}{D}\,[\log b_c - \log b_o] \text{ if } m = 2. \end{cases} \tag{1.5}$$

It is easy to see from (1.5) that while the effect of b_c on L is minimal, the effect of b_o is significant. For instance, an increase from 4 in. to 5 in. in b_c increases L only by $\dfrac{1}{20D}$, when $m = 4$. For the same m, the increase in L when b_o is decreased from 0.01 in. to 0.001 in. is $\dfrac{900}{D}$. Also, the initial crack size is difficult to measure since it is so small. It would therefore be useful to model the distribution b_o. We do this in the next section and show how this leads to a simple distribution for L.

To the best of our knowledge, there have been only two previous attempts made to introduce probability models in the problem of fatigue crack growth. Racicot[3] assumed fixed values for b_o and b_c, and various probability distributions for the several factors in D. He then obtained empirical results for the distribution of L by simulation on a computer. Hanagud and Uppaluri[2] used a formula for crack growth rate which is slightly more complex than (1.1) and involves a quantity r which is the ratio of the maximum stress intensity factor to the minimum stress intensity factor. They assumed that both ΔK and r were random, and obtained approximations to the expected crack size growth rate. They used these results to study improvements in reliability obtainable by examination and repair. Since examination and repair are not possible in a gun barrel operating in the field, their results are not applicable to our problem.

2. THE PROBABILITY MODEL

The first few rounds of firing produce a heat check pattern. Let this pattern contain N cracks having sizes $C_1, \ldots, C_N$. Then the initial crack size b_o is given by

$$b_o = \max \{C_1,\ldots,C_N\}.$$

In our probability model, we assume the $C_1, C_2,\ldots$ are independent and identically distributed according to a uniform distribution on [0,B]. We also assume that N is an independent variable governed by a Poisson distribution with parameter λ. A physical interpretation for B and λ is as follows: B represents the maximum possible initial crack size and λ is a measure of the number of cracks in the heat check pattern. Our probability model depends on just these two easily interpretable quantities. Thus,

$$\begin{aligned} P(b_o \leqslant b) &= P(\max(C_1,\ldots,C_N) \leqslant b \mid N \geqslant 1) \\ &= \frac{\sum_{n=1}^{\infty} P(\max(C_1,\ldots,C_n) \leqslant b)\, P(N = n)}{\sum_{n=1}^{\infty} P(N = n)} \\ &= \frac{\sum_{n=1}^{\infty} \left(\frac{b}{B}\right)^n e^{-\lambda}\lambda^n/n!}{\sum_{n=1}^{\infty} e^{-\lambda}\lambda^n/n!}, \end{aligned}$$

so that

$$P(b_o \leqslant b) = \frac{e^{\lambda b/B}-1}{e^{\lambda}-1}, \quad 0 \leqslant b \leqslant B. \tag{2.1}$$

The distribution in (2.1) is our probability model for the initial crack size. It is physically motivated and simple. A sample application of this model given at the end of this section provides more support for the model.

If the distribution of D (for definition, see (1.4)) or the distribution of the factors that enter D are known, one can combine that knowledge with the probability distribution of b_o in (2.1) to obtain the distribution of the life L of the barrel.

As a sample application, we will obtain an approximation to the distribution of L. Assume that D is known and $m = 4$. Formula (1.5) states that

$$L = \frac{1}{Db_o} - \frac{1}{Db_c}.$$

Thus, for any $x > 0$,

$$\begin{aligned} P\left\{L \leqslant \left(\frac{1}{DB} = \frac{1}{Db_c}\right) + \frac{x}{\lambda DB}\right\} \\ = P\left\{\frac{b_o}{B} \geqslant (1 + \frac{x}{\lambda})^{-1}\right\} \end{aligned}$$

$$\doteq P\left\{\frac{b_o}{B} \geqslant 1 - \frac{x}{\lambda}\right\}$$

$$= \frac{e^{\lambda} - e^{\lambda(1-x/\lambda)}}{e^{\lambda} - 1} \quad \text{from (2.1)}$$

$$\doteq 1 - e^{-x}, \tag{2.2}$$

where $\doteq$ stands for approximately equal to and we have used the relations

$$\left[1 + \frac{x}{\lambda}\right]^{-1} \doteq 1 - \frac{x}{\lambda} \text{ and } e^{\lambda} - 1 \doteq e^{\lambda},$$

which are good approximations even for moderate values of λ. The conclusion in (2.2) may be summarized as follows:

- If the number of all crack sizes in the heat check is moderately large, the life of a gun barrel has an exponential distribution with a location shift.
- In the general case, where D is assumed to be random, the distribution of the life of the barrel becomes a mixture of shifted exponential distributions.

The assumptions of a uniform distribution for initial crack size and a Poisson distribution for the number of cracks in the heat check pattern are not required for a result like (2.2) and were made here just to illustrate the proof. In fact, it can be shown from the theory of limiting distributions for the maximum (or minimum) of large numbers of random variables that the distribution of L can be approximated by a shifted Weibull distribution. In the general case where D is assumed to be random, the distribution of L can be approximated by a mixture of shifted Weibull distributions with a common shape parameter.

It is hoped that currently available data on gun barrels will be studied further to test the validity of our model.

3. CONCLUSIONS

(1) A simple probability model for the initial crack size is given in (2.1).

(2) It is derived, as a consequence, that the life of a barrel has a shifted exponential distribution.

(3) Under practically no assumptions, it can be shown that the life of a barrel has a shifted Weibull distribution.

(4) When D as defined in (1.4) is assumed to be random, the distribution of the life of a barrel is a mixture of shifted Weibull distributions with a common shape parameter.

REFERENCES

[1] Davidson, Thomas E., Joseph F. Throop, and Albert E. Reiner, "The Role of Fracture Toughness and Residual Stresses in the Fatigue Fracture Behavior of Large Thick-Walled Pressure Vessels," National Conference on Fluid Power, *26*, 102-114, (1972).

[2] Hanagud, S., and B. Uppaluri, "A Reliability-Based Cost Effective Fail-Safe Design Procedure, Proc. Seventeenth Structures SDM Meeting A.I.A.A. 537-548 (1976).

[3] Racicot, R. L., Private communication, (1976).

[4] Throop, Joseph F., "Fatigue Crack Growth in Thick-Walled Cylinders," National Conference on Fluid Power, *26*, 115-131, (1972).

MULTIPLE-ATTRIBUTE DECISION MAKING WITH PARTIAL INFORMATION: THE COMPARATIVE HYPERVOLUME CRITERION

Johnnie R. Charnetski

College of Administration and Business
Louisiana Tech University
Ruston, Louisiana

Richard M. Soland

Departement de Génie Industriel
Ecole Polytechnique de Montréal
Montréal, Quebec, Canada

ABSTRACT

A new approach is presented for analyzing multiple-attribute decision problems in which the set of actions is finite and the utility function is additive. The problem can be resolved if the decision makers (or group of decision makers) specifies a set of nonnegative weights for the various attributes or criteria, but we here assume that the decision maker(s) cannot provide a numerical value for each such weight. Ordinal information about these weights is therefore obtained from the decision maker(s), and this information is translated into a set of linear constraints which restrict the values of the weights. These constraints are then used to construct a polytope W of feasible weight vectors, and the subsets H_i (polytopes) of W over which each action a_i has the greatest utility are determined. With the Comparative Hypervolume Criterion we calculate for each action the ratio of the hypervolume of H_i to the hypervolume of W and suggest the choice of an action with the largest such ratio. Justification of this choice criterion is given, and a computational method for accurately approximating the hypervolume ratios is described. A simple example is provided to evaluate the efficiency of a computer code developed to implement the method.

INTRODUCTION

We consider finite action decision problems of the following nature: The decision maker(s) must choose one action from a finite set A of feasible actions. Each action $a_i \in A$, $i = 1, \ldots, m$, is evaluated with respect to a finite set $C = \{c_j | j = 1, \ldots, n\}$ of judgment criteria (attributes). Let s_{ij} be the raw score of action a_i with respect to attribute c_j; these scores may be on either ordinal or interval scales.

We suppose an action will be chosen on the basis of maximum utility, where the utility of action a_i is $u^*(a_i) = u(s_{i1}, \ldots, s_{in})$. We further assume that the utility function u is additive (see Fishburn [5-7]) so that

$$u^*(a_i) = u(s_{i1}, \ldots, s_{in}) = \sum_{j=1}^{n} u_j(s_{ij}). \tag{1}$$

It then follows from the basic results of additive utility theory that for the purposes of decision making we may replace the functions u_j of the raw scores s_{ij} by $w_j v_{ij}$, where v_{ij} is a relative value of the raw score s_{ij} based on the set of all possible scores s_{kj} with respect to attribute c_j, and each w_j is a positive weight. Furthermore, without loss of generality, we may choose the relative values so that for each attribute c_j, the most desired raw score s_{i^*j} of the m scores is mapped into the value $v_{i^*j} = 10$ while the least desired raw score $s_{i'j}$ of the m scores is mapped into the value $v_{i'j} = 0$. (This assumes that not all the values $u_j(s_{ij})$, $i = 1, \ldots, m$, are the same. If this were so, there would be no need to consider attribute c_j.) All the other raw scores s_{ij} for the various actions a_i are mapped into values between 0 and 10 inclusive. (Any finite values a and b, with $a < b$, could be used instead of 0 and 10, respectively; we choose the interval [0, 10] for its simplicity and intuitive appeal.) Also without loss of generality, we may choose the weights w_j so that $\sum_j w_j = 1$.

Thus we write the utility of each action a_i as

$$u^*(a_i) = \sum_{j=1}^{n} w_j v_{ij}. \tag{2}$$

Since the v_{ij} depend on the raw scores s_{ij} and the preferences of the decision maker(s) for these scores, we may asume that he (they) can provide the appropriate v_{ij} values without undue difficulty. We therefore define the m by n matrix V of the v_{ij} values: for each column j the values v_{ij} lie between 0 and 10 inclusive, with at least one of them being 0 and at least one of them being 10.

If we define the n by 1 column vector w of weights and let V_i be the ith row of V we may rewrite (2) as

$$u^*(a_i) = V_i w. \tag{3}$$

If the decision maker(s) provide(s) an appropriate weight vector $w = w^*$, then (2) or (3) should be used as the basis for selecting an action. In many cases, however, the decision maker(s) may not be willing to provide a particular w. In the case of an individual decision maker this may be simply due to the fact that he or she cannot articulate his or her preferences with such precision. Indeed, it may be somewhat unusual to find a decision maker who will specify a particular w. In the case of a group of decision makers, there may be considerable disagreement about the appropriate weight vector w. This could be the situation if, for example, they represent different sectors of society and the decision problem encompasses different criteria which are economic, environmental, and political in nature.

We may therefore distinguish two extreme situations that can apply to the state of knowledge of the decision maker(s) with respect to w:

(i) Nothing is known about w except that it lies in the set

$$\left\{ w \in E^n \middle| \sum_{j=1}^{n} w_j = 1, \text{ all } w_j \geqq 0 \right\}.$$

(ii) It is known that $w = w^*$.

We term case (i) multiple-attribute decision making with *no information* and case (ii) multiple-attribute decision making with *complete information*, and respectively refer to a multiple-attribute

problem with no information or complete information. We wish to consider the general case between these two extremes, and we term it multiple-attribute decision making with *partial information.* It will include the two cases above as specific extremes.

We therefore define the set W as the set of weight vectors that the decision maker(s) deem(s) feasible, i.e., may be an appropriate one in light of his subjective feelings (or in light of their range of agreement if there are several decision makers). In case (i) above, therefore, we have

$$W = \left\{ w \in E^n \middle| \sum_{j=1}^{n} w_j = 1, \text{ all } w_j \geqq 0 \right\} \equiv \overline{W},$$

while in case (ii) we have

$$W = \{w^*\}.$$

Having thus introduced the notion of multiple-attribute decision making with partial information, we shall, in the next section, discuss the construction and characterization of W. After that we will introduce the Comparative Hypervolume Criterion for choosing an action when the set W is not a unique point. The following section details the Monte Carlo method that has been developed to accurately approximate the numerical quantities used by the Comparative Hypervolume Criterion. The final section summarizes the numerical method and presents an illustrative example that has been used to partially evaluate the accuracy and efficiency of the Monte Carlo method and the computer program implementing it.

We note here that a similar approach has been presented in [3] for the case of no information and a minor extension of it. The numerical approach of [3] is only efficient for the case of no information, however, and cannot be used in the general case of partial information. The Monte Carlo technique presented herein is quite different and much more efficient.

THE SET W OF FEASIBLE WEIGHT VECTORS

We have seen above that the set W of feasible weight vectors is $\overline{W}$ in the case of no information and the singleton set $\{w^*\}$ in the case of complete information. In the general case we may suppose that W is defined by equality and inequality constraints involving the components $w_1, \ldots, w_n$ of w. In the present situation, however, we will confine ourselves to *linear* constraints involving the weights. These may arise very naturally as follows.

We can translate the statement "attribute c_j is at least as important as attribute c_k" on the part of the decision maker(s) into the mathematical statement $w_j \geqq w_k$. As a direct extension of this that is compatible with the foundations of additive utility [5], we let J and K be subsets of $N \equiv \{1, \ldots, n\}$ and translate the statement "the set of attributes with indices in J, taken together, are at least as important as the set of attributes with indices in K" into the mathematical statement

$$\sum_{j \in J} w_j \geqq \sum_{k \in K} w_k.$$

With similar reasoning, constraints of the form

$$\sum_{j \in J} w_j \leqq b_J,$$

where b_J is a constant between zero and one, may be elicited from the feelings and statements of the decision maker(s). The important point is that the decision maker(s) agree(s) that weight vectors not satisfying such constraints are not to be considered.

In general, then, we shall assume that the partial information possessed by the decision maker(s) concerning w has been encoded in a set of linear constraints on $w_1, \ldots, w_n$. To these we add the normalizing constraint $\sum_j w_j = 1$. Without loss of generality we can convert all the constraints to inequality constraints; hence we may write W as

$$W = \{w \in E^n \mid Aw \leqq b,\ w \geqq 0\}, \tag{4}$$

where A is the s by n matrix of constraint coefficients and b is the s by 1 right-hand side vector. Thus W is a bounded polytope in E^n. $W \subset \overline{W}$, and in the case of no information, $W = \overline{W}$. In the case of complete information W is a single point.

THE COMPARATIVE HYPERVOLUME CRITERION

Now that we have constructed W, how shall we use it in aiding the decision maker(s) to select an action a_i to implement? For every $w \in W$ we can compute, for each $a_i \in A$, the utility value $u^*(a_i) = V_i w$, and hence find the action or actions with highest utility value for this particular w. Let us call such an action an optimal one with respect to w and denote by $A(w)$ the set of such optimal actions with respect to w. Clearly, if there is an action a_i such that $a_i \in A(w)$ for all $w \in W$, then we can confidently claim that this a_i is the one to choose. Given A, b, and V, we can check, but not necessarily easily, for the existence of such a clearly optimal a_i (a necessary and sufficient condition for $a_i \in A(w)$ for all $w \in W$ is that $a_i \in A(w^k)$ for all extreme points w^k of W). But it will be rare indeed that such a clearly optimal action exists, so we must search further for a basis for choosing an action.

If there is an action a_i such that $a_i \in A(w)$ for "most" $w \in W$ (but not all $w \in W$), then it seems reasonably wise to select this action; it is the one most "likely" to lead to the highest utility of any of the actions in A. Here, "likely" is taken with respect to a uniform probability distribution of the $w \in W$. What we are now suggesting, therefore, is to measure for each $a_i \in A$ the relative frequency with which $a_i \in A(w)$. We then suggest as optimal an action with the largest such relative frequency. Stated another way, we define an optimal action to be one which maximizes the probability of yielding a utility value at least as high as that of all other actions for a randomly selected $w \in W$. For reasons which will shortly become apparent, we call this criterion the Comparative Hypervolume Criterion (CHC). Besides its intuitive appeal as illustrated above, its compatibility with Bernoullian utility constructs has been demonstrated in Ref. [1].

Let us now be more exact about the CHC. Define, for $i = 1, \ldots, m$,

$$H_i = \{w \in W \mid a_i \in A(w)\}, \tag{5}$$

where, to repeat,

$$A(w) = \{a_i \in A \mid V_i w \geqq V_k w,\ k = 1, \ldots, m\}. \tag{6}$$

Thus we have the equivalent characterization

$$H_i = \{w \in W \mid V_i w \geqq V_k w,\ k = 1, \ldots, m\}. \tag{7}$$

We then measure the relative frequency with which $a_i \in A(w)$ by

$$r_i \equiv \int_{H_i} dw \Big/ \int_W dw, \tag{8}$$

where the element of integration dw in (8) depends on the dimensionality of W (which will be discussed below). Thus, r_i is a ratio of hypervolumes. The Comparative Hypervolume Criterion then suggests the choice of an action a_i such that $r_i \geqq r_k$, $k = 1, \ldots, m$.

The computational difficulty faced by the CHC is the computation of the ratios $r_1, \ldots, r_m$. Although W has an explicit characterization as a polytope and the same is true of H_i by virture of (7), it is very difficult to carry out the integrations required by (8). We have developed a Monte Carlo simulation method to accurately approximate the ratios $r_1, \ldots, r_m$ needed by the CHC; this is detailed in the next section.

THE MONTE CARLO SIMULATION METHOD

In Ref. [3] is given a relatively efficient Monte Carlo method for evaluating the hypervolume ratios in the case of multiple-attribute decision making with no information. In this case $W = \overline{W}$, so a random point w within W can be easily generated by drawing n uniform random numbers and then normalizing them to sum to one. The quantities $V_i w$, $i = 1, \ldots, m$ are compu ed and compared, and a record is kept of the value or values of i that maximize(s) $V_i w$. If k_i s the count of the number of drawings for which action a_i yields the maximum value of $V_i w$ in a total sample size of g, then k_i/g is the appropriate estimate of r_i.

This approach might be extended to the case of partial information by discarding the point w drawn if $w \in W$, but this sampling will be highly inefficient if W is small relative to $\overline{W}$. Moreover, it will not work at all if the dimension of W (the number of independent vectors needed to span it) is less than $n-1$ (as will be the case if W is partially defined by an equality constraint other than $\sum w_i = 1$).

In Ref. [2] we outlined a general Monte Carlo approach for computing statistical measures for a linear function defined over a polytope. That general approach will be adapted and extended here in order to yield an efficient numerical procedure for approximating the hypervolume ratios. The basic idea is to change coordinate systems, taking a vertex of W as the new origin. The new coordinate system uses the ordinary Euclidean distance ρ and a set of angular coordinates. For a specific set of angular coordinate values, the integration with respect to ρ required in the numerator and denominator of (8) may be carried out explicitly. Integration with respect to the angular coordinates cannot be explicitly carried out, however, and is instead replaced by a Monte Carlo drawing of the possible sets of angular coordinates. These correspond to vectors randomly directed into W from the vertex chosen. As will be seen, the method requires the vertex chosen to be a nondegenerate one. Now we present the method in detail.

Let d be the dimension of $W (d \leq n-1)$ and let γ be a nondegenerate vertex or extreme point of W. Thus, γ is formed by the intersection of exactly d linearly independent hyperplanes from the set of hyperplanes defining W. We choose γ as the origin in a spherical coordinate system utilizing the Euclidian distance ρ and $d-1$ angles. We will use ϕ to designate the vector of angles. For an arbitrary point $w \in W$ we have

$$\rho = ||w - \gamma|| = \left[\sum_{j=1}^{n} (w_j - \gamma_j)^2\right]^{1/2},$$

and the information needed to compute the appropriate value of ϕ for w (assuming $w \neq \gamma$) is contained in the unit vector $\delta = (w - \gamma)/||w - \gamma||$. Thus the denominator in (8), call it $S(W)$, may be expressed as

$$S(W) = J \int\int_{W} \rho^{d-1} \, d\rho d\phi,$$

where J is a constant and $J\rho^{d-1}$ is the Jacobian of the coordinate transformation. An analogous expression, call it $S(H_i)$, holds for the numerator of (8) so that we have

$$r_i = S(H_i)/S(W) = \iint_{H_i} \rho^{d-1}\, d\rho d\phi / \iint_{W} \rho^{d-1}\, d\rho d\phi. \tag{9}$$

We shall shortly show that for any specific ϕ we can explicitly perform the indicated integration with respect to ρ. We cannot perform the integration with respect to ϕ, however, and instead will use Monte Carlo simulation to generate random values of ϕ.

With ϕ defined as above for any $w \in W$ (take $\delta = 0$ for $w = \gamma$), we may write $w \in W$ as $w = \gamma + \rho\delta$. Then define Δ as the set of all such unit vectors directed from γ into W, i.e.,

$$\Delta = \{\delta \mid \delta = (w - \gamma)/||w - \gamma||,\ w \in W,\ w \neq \gamma\}. \tag{10}$$

Thus for every ϕ such that $w = (\rho, \phi) \in W$ we can equivalently describe w as $\gamma + \rho\delta$ since there is a unique correspondence between ϕ and δ. Now rewrite $S(H_i)$ as

$$S(H_i) = \iint_{W} \Omega(H_i)\, \rho^{d-1}\, d\rho d\phi,$$

where $\Omega(H_i)$ is the indicator function of the set H_i. Now that both $S(H_i)$ and $S(W)$ are expressed in terms of integration over W, we will explicitly perform the integrations with respect to ρ.

Given an arbitrary $\delta \in \Delta$, denote by $\rho(\delta)$ the largest value of ρ such that $\gamma + \rho\delta \in W$. Thus, by the convexity of W, the limits of integration on ρ in $S(H_i)$ and $S(W)$ are 0 and $\rho(\delta)$, respectively. We can find $\rho(\delta)$ as follows: Letting A_l be the lth row of the s by n matrix A partially defining W through $Aw \leqq b$, $\rho(\delta)$ is the smallest distance to a plane $A_l w = b_l$ such that $A_l\gamma < b_l$. Of course $A_l(\gamma + \rho\delta) = b_l$ cannot occur for positive ρ if $A_l\delta \leqq 0$, so we must exclude this case. We thus find

$$\rho(\delta) = \min_l \{(b_l - A_l\gamma)/(A_l\delta) \mid A_l\gamma < b_l,\ A_l\delta > 0\}. \tag{11}$$

In the direction δ, the indicator function $\Omega(H_i)$ is 1 for ρ in the interval $[\underline{h}_i(\delta), \bar{h}_i(\delta)]$ and 0 for ρ outside this interval, where $\underline{h}_i(\delta)$ and $\bar{h}_i(\delta)$ are the smallest and largest nonnegative values of ρ, for $\rho \leqq \rho(\delta)$, respectively, such that $V_i(\gamma + \rho\delta) \geqq V_k(\gamma + \rho\delta)$, $k = 1, \ldots, m$. (If no nonnegative value of $\rho \leqq \rho(\delta)$ exists which satisfies this inequality, we take $\underline{h}_i(\delta) = \bar{h}_i(\delta) = 0$.) We will see below that it is a simple matter to determine all the intervals $[\underline{h}_i(\delta), \bar{h}_i(\delta)]$.

Performing the integrations with respect to ρ in (9) now, we obtain

$$r_i = \frac{\int \left[\left(\bar{h}_i(\delta)\right)^d - \left(\underline{h}_i(\delta)\right)^d \right] d\phi}{\int \left[\rho(\delta)\right]^d d\phi}, \tag{12}$$

where the integrations with respect to ϕ must be performed over all ϕ such that the corresponding δ lies in Δ. We cannot perform these integrations easily, so we shall instead approximate (12) by randomly generating direction vectors $\delta \in \Delta$. For each such δ^k so generated we can easily compute the quantities $\rho(\delta^k)$, $\underline{h}_i(\delta^k)$, and $\bar{h}_i(\delta^k)$, $i = 1, \ldots, m$. With a sample of direction vectors δ of size g, we then approximate (12) by

$$\bar{r}_i = \frac{\sum_{k=1}^{g} \left[\left(\bar{h}_i(\delta^k)\right)^d - \left(\underline{h}_i(\delta^k)\right)^d \right]}{\sum_{k=1}^{g} \left[\rho(\delta^k)\right]^d}. \tag{13}$$

Equation (13) represents a sampling formulation of the hypervolume ratio r_i of (8); it is a statistical approximation of r_i with a sampling error that is a function of the sample size g. Thus, for g sufficiently large, all $\bar{r}_i$ will approximate their respective r_i accurately enough to be used in their place in connection with the CHC.

We have thus far failed to treat three important points: (a) determination of the dimension d of W, (b) explicit computation of the $\underline{h}_i(\delta)$ and $\bar{h}_i(\delta)$, and (c) generation of random direction vectors $\delta \in \Delta$. The last of these depends on the fact that γ is a nondegenerate extreme point of W, and the first is aided considerably by this fact. Being nondegenerate, γ is formed by the intersection of exactly d linearly independent hyperplanes from the set of hyperplanes defining W. As demonstrated elsewhere [1], the edges of W at γ, i.e., the intersection of the d linearly independent hyperplanes taken $(d\text{-}1)$ at a time, comprise a set of d linearly independent vectors along which lie the extreme points of W adjacent to γ. Thus there are exactly d adjacent extreme points, and since γ is nondegenerate (in the linear programming sense also), each of these adjacent extreme points can be generated by part of one simplex pivot. Denote these adjacent extreme points by $\gamma^1, \ldots, \gamma^d$. Note that the number of them serves to define the dimension d, thus taking care of point (a) above.

Define the unit vectors $q^1, \ldots, q^d$ by

$$q^l = (\gamma^l - \gamma)/||\gamma^l - \gamma||, \quad l = 1, \ldots, d. \tag{14}$$

This set of linearly independent direction vectors spans Δ and will be used to generate random vectors $\delta \in \Delta$. Let q be a convex combination of $q^1, \ldots, q^d$, i.e., $q = \sum_{l=1}^{d} t_l q^l$, where $\sum_{l=1}^{d} t_l = 1$ and all $t_l \geqq 0$. Thus $q/||q||$ is an element of Δ, and in fact we have

$$\Delta = \{q/||q|| \mid q = \sum_{l=1}^{d} t_l q^l, \ \sum_{l=1}^{d} t_l = 1, \ \text{all } t_l \geqq 0\}. \tag{15}$$

Since $q^1, \ldots, q^d$ are linearly independent, each unique convex combination of them generates a unique vector q and hence a unique vector $\delta = q/||q||$ which is an element of Δ. Thus, in order to generate a random element of Δ we might simply generate d uniform random numbers, normalize them to form a random convex combination, and then use them in expression (15) to obtain a random element $\delta \in \Delta$. Unfortunately, this straightforward approach leads to a biased choice of $\delta \in \Delta$. It is not difficult to show, however, that if all pairs of unit vectors (q^k, q^l), $l \leqslant k < l \leqslant d$, form acute angles, then an unbiased choice results if we use

$$\delta = \sum_{l=1}^{d} t_l q^l / || \sum_{l=1}^{d} t_l q^l ||, \tag{16}$$

where the t_l are independent uniform random numbers on (0, 1) and a sequence $(t_1, \ldots, t_d)$ is discarded if $||\Sigma_l t_l q^l|| > 1$. If, however, some pairs of unit vectors (q^k, q^l) form obtuse angles, then it is necessary to further restrict the sequences $(t_1, \ldots, t_d)$ used to those for which $||\Sigma_l t_l q^l|| \leqslant u$, where $u = (1 - \cos^2\beta)^{1/2}$ and β is the largest angle between pairs of unit vectors (q^k, q^l). This takes care of point (c).

Now we deal with point (b). Given the nondegenerate extreme point γ, a randomly generated direction vector $\delta \in \Delta$, and $\rho(\delta)$, there exists at least one a_k such that $V_k\gamma \geqq V_i\gamma$ for $i = 1, \ldots, m$. Without loss of generality we will assume that a_1 has this property, i.e., is optimal for the weight vector γ. Then $h_1(\delta) = 0$. Now suppose a_1 is the unique optimal action at γ; i.e., $V_1\gamma > V_i\gamma$ for $i > 1$. Then either (a) $V_1(\gamma + \rho\delta) > V_i(\gamma + \rho\delta)$ for $i > 1$ and all $\rho \in [0, \rho(\delta)]$, in which case $\bar{h}_1(\delta) = \rho(\delta)$, or (b) there exists a smallest scalar $z \in (0, \rho(\delta)]$ such that $V_1(\gamma + z\delta) = V_i(\gamma + z\delta)$ for some $i > 1$. In case (a) we take $\underline{h}_k(\delta) = \bar{h}_k(\delta) = 0$ for $k > 1$. In case (b) we clearly have $\bar{h}_1(\delta) = z$, where z is explicitly given by

$$z = \min_{k>1}\{(V_1 - V_k)\gamma/(V_k - V_1)\delta \mid (V_1 - V_k)\gamma/(V_k - V_1)\delta > 0\}. \tag{17}$$

If z from (17) exceeds $\rho(\delta)$, then case (a) applies. Suppose case (b) applies and $i = 2$. Then we have $\underline{h}_2(\delta) = z = \bar{h}_1(\delta)$, and the above method for finding $\bar{h}_1(\delta)$ can now be used to find $\bar{h}_2(\delta)$. In a straightforward and efficient way, therefore, all the intervals $\left[\underline{h}_i(\delta),\ \bar{h}_i(\delta)\right]$, $i = 1, \ldots, m$, may be determined.

We have thus far ignored the case in which a_1 is *not* the unique optimal action at γ, i.e., in which $V_1\gamma = V_i\gamma$ for at least one $i > 1$. To see whether we want a_1 or one of the other equally good actions at γ we have merely to compare the values of $V_1(\gamma + \epsilon\delta)$ and $V_i(\gamma + \epsilon\delta)$ for a sufficiently small positive ϵ and use the action with highest such value. Note that while this form of tie among several actions may occur at γ, it will not occur in practice at the point $\gamma + \bar{h}_1(\delta)$ because δ is chosen randomly and hence the probability of such a tie is effectively zero.

One more point deserves mention here. The Monte Carlo method requires that γ be a nondegenerate extreme point of W. We can clearly increase the likelihood of finding such a nondegenerate extreme point to use by eliminating obviously redundant constraints from the definition of W and keeping any equality constraints defining W as equalities instead of two inequalities. Despite these precautions, it is conceivable that W may not possess any nondegenerate extreme points, or that a reasonable amount of searching does not identify one. In such a case it is computationally reasonable to modify a degenerate extreme point by making the basic variables which are zero slightly positive. W will thus be modified and will have at least one new extreme point, a nondegenerate one. The Monte Carlo method can then be used, and, since this type of perturbation can be made arbitrarily small, the errors introduced by it can also be made arbitrarily small.

We have now treated all aspects of the Monte Carlo method developed to approximate the hypervolume ratios $r_1, \ldots, r_m$ by $\bar{r}_1, \ldots, \bar{r}_m$. In the next section we will give a concise step-by-step summary of the method and then present an illustrative example whose obvious solution serves as a check on the Monte Carlo method and the computer program written to implement it.

COMPUTATIONAL SUMMARY AND AN EXAMPLE

The computational method for computing $\bar{r}_1, \ldots, \bar{r}_m$ by (13) may be summarized as follows:

(i) Select a nondegenerate vertex of W, and label it γ. This can be accomplished by generating an arbitrary linear objective function to be maximized over W, and, utilizing the simplex method, continuing to pivot until a suitable vertex is obtained.

(ii) Determine the vertices $\gamma^1, \ldots, \gamma^d$ adjacent to γ and then the unit vectors $q^1, \ldots, q^d$. Again using the simplex method, each γ^l is obtained by pivoting into the basis represented by γ an appropriate nonbasic variable. Slack variables must be considered too. The dimension of W, d, is determined by the number of adjacent vertices.

For $k = 1, \ldots, g$, carry out steps (iii) through (vi) (in which the dependence on k is suppressed).

(iii) Generate sets of d uniform random numbers $(t_1, \ldots, t_d)$ until a set is found such that $||q|| \leqslant 1$ or $||q|| \leqslant u$ (see discussion following (16)), where $q = \Sigma_l t_l q^l$. Then generate the direction vector $\delta = q/||q||$.

(iv) Determine the upper integration limit $\rho(\delta)$ from (11).

(v) Determine the intervals $[h_i(\delta), \bar{h}_i(\delta)]$, $i = 1, \cdots, m$.

(vi) Compute the kth terms in the summations in the numerator and denominator of (13) and add them to the previous subtotals.

(vii) Compute all $\bar{r}_i$ via (13).

Once the $\bar{r}_i$ are available as accurate approximations of the r_i, the Comparative Hypervolume Criterion suggests the choice of the action a_i with largest value of $\bar{r}_i$.

A computer code was written in FORTRAN to implement the numerical method developed above, and the following simple example was used to get some idea of the code's accuracy and running time. The example had $i = 3$ and $n = 52$. V_i (of size 52) had a 10 in column i and a 0 in the other 51 columns, $i = 1, 2, 3$. $W = \overline{W}$ was used. By virtue of the symmetry present it is easily seen that $r_1 = r_2 = r_3 = 1/3$.

Table 1 shows the numerical values obtained for the $\bar{r}_i$ for the first g random vectors δ, $g = 100, 200, 400, 900, 2000$. The results are quite accurate for the larger values of g and are reasonably accurate even for small values of g. The cpu time required for each sample of 100 direction vectors averaged about 1.25 seconds on an IBM 370/158 computer.

TABLE 1. *Computed Hypervolume Ratios for the Example*

Sample Size g	Computed Ratios		
	$\bar{r}_1$	$\bar{r}_2$	$\bar{r}_3$
100	0.2821	0.3883	0.3295
200	0.3244	0.3424	0.3332
400	0.3362	0.3299	0.3340
900	0.3230	0.3449	0.3321
2000	0.3343	0.3352	0.3305

DISCUSSION

The Comparative Hypervolume Criterion has been presented as a rationale and method by which a decision maker or group of decision makers can select one action from a finite set of actions when the selection must account for multiple criteria or attributes. It does not require that the decision maker(s) specify exact values by which to weight the various attributes, and this could be an important consideration in practical decision making situations. If the comparative hypervolume ratio r_i for a particular action a_i is significantly greater than one-half, then the decision maker(s) should be able to adopt it with confidence. If, on the other hand, there is no action with a fairly large ratio, then more caution is called for. Since r_i is the proportion of weight vectors for which a_i has a utility value at least as high as *all* other actions, it is possible that *no* action will have a very large ratio. This might well be the case, for instance, if several different actions have rows V_i of V which are fairly similar. It would then be appropriate to use the CHC to make comparisons among the actions in specific subsets of A. For example, some judicious pairwise comparisons might enable the decision maker(s) to eliminate a number of actions. The CHC might then point rather definitively to one of the remaining ones.

Thus, even if it is not used as the sole basis for selecting an action, use of the CHC can greatly aid the decision maker(s) in excluding some possible choices or in evaluating the effect of various constraints placed on the attribute weights. To aid decision makers in these ways, the computer code could easily be operated in an interactive mode.

BIBLIOGRAPHY

[1] Charnetski, J. R., "The Multiple Attribute Problem with Partial Information: The Expected Value and Comparative Hypervolume Methods," Ph.D. Thesis, The University of Texas at Austin, Austin, Texas (August 1973).

[2] Charnetski, J. R., and R. M. Soland, "Statistical Measures for Linear Functions on Polytopes," Operations Research *24*, 201-204 (1976).

[3] Connelly, J. J., C. A. Allen, R. L. Holmes, and R. M. Soland, "South Florida Regional Airport Study: Tasks K-1 through K-4," Research Analysis Corporation, Client Report RAC-CR-63, McLean, Virginia (July 1972).

[4] Davidson, F., "Dimensions of Utility in a Regional Planning Context," Decis. Sci. *5*, 91-101 (1974).

[5] Fishburn, P. C., "A Note on Recent Developments in Additive Utility Theories for Multiple-Factor Situations," Operations Research *14*, 1143-1148 (1966).

[6] Fishburn, P. C., "Additive Utilities with Finite Sets: Applications in the Management Sciences," Naval Research Logistics Quarterly *14*, 1-10 (1967).

[7] Fishburn, P. C., *Utility Theory for Decision Making*, (John Wiley & Sons, New York, 1970).

[8] Fishburn, P. C., "Lexicographic Orders, Utilities and Decision Rules: A Survey," Working Paper, Pennsylvania State University, (1972).

[9] MacCrimmon, K. R., "An Overview of Multiple Objective Decision Making," in *Multiple Criteria Decision Making*, J. L. Cochrane and M. Zeleny (eds.), pp. 18-44, (University of South Carolina Press, Columbia, S. C., 1973).

AN ALGORITHM FOR A CLASS OF LOADING PROBLEMS

Ming S. Hung

College of Business Administration
Cleveland State University
Cleveland, Ohio

J. Randall Brown

College of Business Administration
Kent State University
Kent, Ohio

ABSTRACT

The loading problem we consider is to assign a set of discrete objects, each having a weight, to a set of boxes, each of which has a capacity limit, in such a way that every object is assigned to a box and the number of boxes used is minimized.

A characterization of the assignments is offered and used to develop a set of rules for generating nonredundant assignments. The rules are incorporated into an implicit enumeration algorithm. The algorithm is tested against a very good heuristic. Computational experience shows that the algorithm is highly efficient, solving problems of up to 3600 0-1 variables in a CPU second.

1. INTRODUCTION

One form of the loading problem [2] is to assign a set of m objects, each having a certain weight w_i, $i = 1, \cdots, m$, to a number of boxes, each of which has a weight capacity limit c_j, $j = 1, \cdots, n$, in such a way that every object is assigned to a box and the number of boxes used is minimized.

There are many applications of the loading problem. Eilon, et al. [3] considered the problem of determining the number of vehicles to carry a given consignment to a destination. It can be used to cut rectangles from larger rectangular sheets [5] and to schedule jobs of given duration on parallel machines [7]. Johnson [10] called it a "bin packing" problem for assigning data files of given lengths to disc tracks.

Let

$x_{ij} = 1$ if object i is assigned box j, 0 otherwise.

$y_j = 1$ if box j is used, 0 otherwise.

The loading problem can be written as the 0-1 linear program:

(p) $$\text{Minimize } z = \sum_{j=1}^{n} y_j$$

subject to

(1) $$\sum_{j=1}^{n} x_{ij} = 1 \qquad i = 1, \cdots, m,$$

(2) $$\sum_{i=1}^{m} w_i x_{ij} \leqslant c_j y_j \qquad j = 1, \cdots, n,$$

(3) $$x_{ij} = 0 \text{ or } 1 \text{ for all } i, j.$$

Other forms of the loading problem have been suggested by Eilon and Christofides [2]. One form is to maximize the total value of the objects that can be assigned to boxes. This has been investigated by Ingargiola and Korsh [9]. Another form is to minimize the unused space in the boxes used. If all box sizes are equal, then this problem is equivalent to our problem (p).

A solution to the loading problem (p) entails two interrelated decisions: which boxes to use and how to assign the objects to the boxes. The second decision is more difficult because of the massive number of combinations that need to be investigated. This paper offers a characterization of such assignments. The characterization has been shown to be very effective in reducing the effort of searching for an optimal solution to the loading problem.

With respect to other solution methods for the loading problem, Eilon and Christofides [2] proposed an implicit enumeration algorithm for problems in which the box capacities are all equal. They also proposed a heuristic which, according to the computational experience conducted for this research, is very good and efficient. A more detailed discussion of the heuristic is given in Section 4. Johnson [10] developed four heuristic algorithms for the bin-packing problem. The heuristics are similar in nature to that of Eilon and Christofides; therefore, their quality should be comparable.

The algorithm presented in this paper is based upon the characterization of the assignments. Computational experience in Section 4 shows that the algorithm, which guarantees optimal solutions, is competitive with Eilon and Christofides' heuristic in terms of solution time.

2. CHARACTERIZING ASSIGNMENT MATRICES

Let $A = [x_{ij}]$ be an m × n matrix of zeros and ones. The rows of A correspond to objects and the columns correspond to boxes. $x_{ij} = 1$ means object i is assigned to box j. Therefore, no row in A contains more than one 1. A will be called an assignment if every row of A has exactly one 1. If some rows are zero rows (null rows), then A is called a partial assignment. (For clarity, object i is denoted by a_i and box j is denoted by b_j in this section.)

For a clearer understanding of the following development, suppose five objects, $a_1, a_2, \cdots, a_5$ are to be assigned to four boxes, $b_1, \cdots, b_4$. Further suppose there are two assignments as shown in Fig. 1.

Suppose objects a_1, a_2, and a_3 have the same weight, i.e., $w_1 = w_2 = w_3$, and boxes b_3 and b_4 have the same capacity, $c_3 = c_4$. Then A_1 and A_2 are essentially the same assignments.

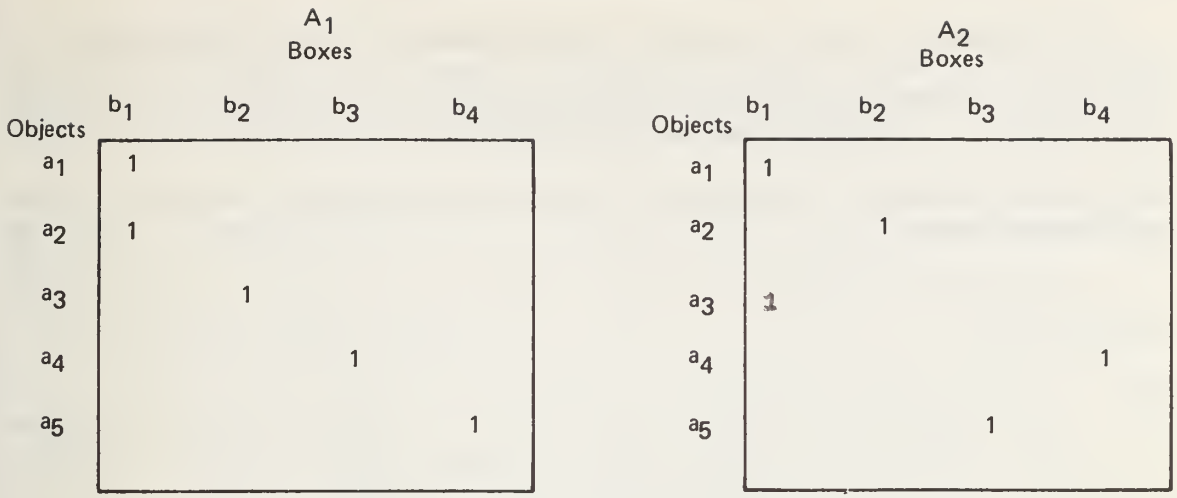


FIGURE 1. Two assignments

First, boxes b_1 and b_2 contain essentially the same objects in both A_1 and A_2 except for the indices of the objects. Second, objects a_4 and a_5 are individually assigned to essentially the same boxes because both boxes have the same capacity. Hereafter we shall refer to two assignments that are essentially the same as "equivalent assignments," and all equivalent assignments but one are "redundant." An assignment that is not equivalent to any other assignment will be called "nonredundant." These definitions are necessary because in an enumeration scheme, optimality of a solution is guaranteed only after all nonredundant solutions are considered, explicitly or implicitly.

To identify the equivalent assignments in general, let us suppose there is an equivalence relation "ρ" (e.g., same weights) among the objects and an equivalence relation "σ" (e.g., same capacities) among the boxes. If a $a_i \rho a_k$, then exchanging rows i and k in the assignment matrix yields an equivalent assignment; so does exchanging columns j and k if $b_j \sigma b_k$. To aid our identification, partition the assignment matrix by blocking together rows equivalent under ρ and columns equivalent under σ. For example, the two matrices in Fig. 1 are blocked and shown in Fig. 2.

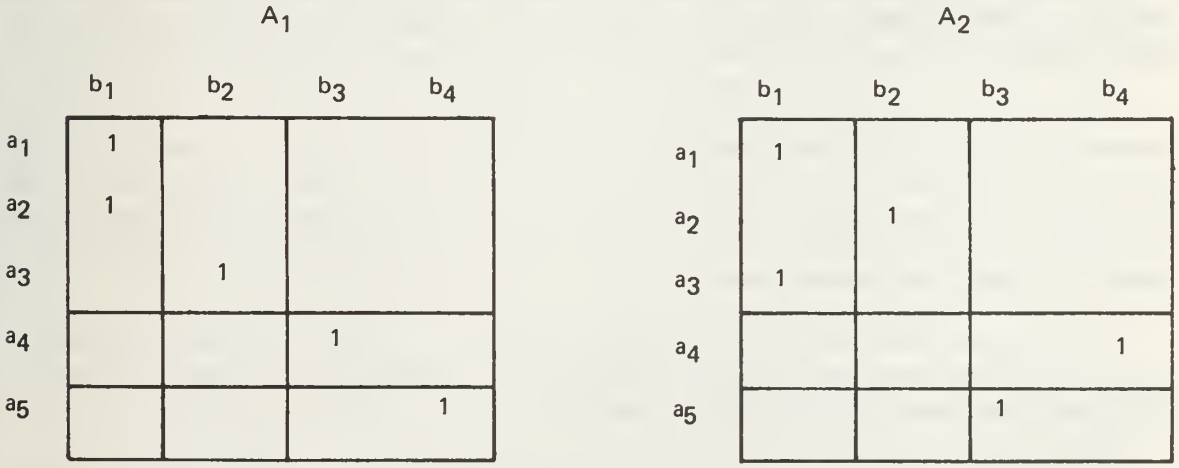


FIGURE 2. Block form of two assignments

An assignment matrix is then transformed into "normal form" by the following steps:

(a) Permute the rows within each row-equivalent block so as to put them in inverse lexicographic order; e.g., (0,1,0,0) precedes (0,0,1,0), which precedes (0,0,0,1).

(b) Permute the columns within each column-equivalent block so as to put them in inverse lexicographic order.

As for our two examples, A_1 is already in normal form, whereas the normal form A_2 is shown in Fig. 3.

A_2

	b_1	b_2	b_4	b_3
a_1	1			
a_3	1			
a_2		1		
a_4			1	
a_5				1

FIGURE 3. Normal form of assignment A_2

It can be seen that A_2 has the same normal form as A_1. Hence, a formal remark:

REMARK: Two assignments are equivalent if their assignment matrices have identical normal forms.

If two partial assignment matrices have the same normal form, then we also assert that for every completion of one assignment, there is an equivalent completion of the other. Thus, in order to avoid generating redundant assingments in an implicit enumeration scheme, one should branch in such a way that every partial assignment is always in normal form.

The normal forms of two equivalent assignments are not necessarily identical because the columns in the same equivalence block may contain an unequal number of 1's among the rows of an equivalence block. For example, assume that b_1 and b_2 in A_1 and A_2 (Fig. 1) are in the same column block and object a_3 is assigned to b_2 in A_2. All other assumptions about A_1 and A_2 remain the same as before. Then the normal form of A_1 and A_2 will be different even though the assignments are equivalent.

The following rules, which were developed by Brown [1], ensure that partial assignments are in normal form and that the normal form is unique. To achieve both purposes, a concept of relatively equivalent columns (boxes) is needed.

DEFINITION: Boxes b_j and b_k (columns j and k) are relatively equivalent if and only if b_j and b_k are equivalent under σ and, in every completely assigned row (object) equivalence block, columns j and k have an equal number of 1's.

A row block is completely assigned if every row has been given a 1. Thus, at the beginning, when no object has been assigned, all equivalent columns (boxes) are also relatively equivalent. Once a row block is completely assigned, some equivalent columns may become relatively unequivalent and remain so for all following partial assignments.

R1: (Selection of branching object.) Object a_i is to be considered for assignment only if (a) object a_{i-1} has been assigned, or (b) i is the smallest index of any object in its equivalence block.

R2: (Selection of branching box.)

1. Object a_i can be assigned to box b_j (i.e., x_{ij} can be fixed to one) if either (a) a_{i-1} was assigned to a box whose index is not greater than j, or (b) i is the smallest index in its equivalence row block. (This rule preserves the lexicographic ordering of rows.)

2. Object a_i can be assigned to box b_j if either (a) column $j-1$ of A is not empty, or (b) j is the smallest index in its relative column equivalence block. (This rule preserves the lexicographic ordering of columns.)

R3: Object a_i can be assigned to box b_j only if the sum of the elements of the row block of a_i in relatively equivalent column $j-1$ is at least one greater than that in column j. (This rule ensures the uniqueness of the normal form and is to be used in conjunction with R1 and R2.)

The above characterization of assignment matrices and the rules for avoiding redundant assignments can be used for any problem involving allocation of a set of discrete entities to another set. Generalized assignment problems [11] and 0-1 multiple knapsack problems [9] belong to this category.

3. THE ALGORITHM

The algorithm presented here is basically an implicit enumeration method whose branching strategy is based on Rules R1-R3. The bounding aspect of the algorithm is adapted from a lower bound suggested by Eilon and Christofides [2].

The lower bound on the objective function value of the loading problem is the fewest possible number of boxes that can contain all the objects. Therefore, it can be computed as follows, where the boxes are indexed in descending order of their capacities:

$$\underline{z}_1 = \min\left\{k : \sum_{j=1}^{k} \geqslant \sum_{i=1}^{m} w_i\right\}. \tag{4}$$

$\underline{z}_1$, the initial lower bound, provides a termination rule for the solutions. If any solution uses $\underline{z}_1$ boxes then the solution procedure is stopped. This bound was found to be very effective in the extensive computational experience described in Section 4.

Before the start of the algorithm, the boxes are indexed in descending order of their capacities and the objects are indexed in descending order of their weights. Eilon and Christofides [2] used an ascending order for the boxes but our experience showed that this often leads to inferior solutions. On the ordering of objects, Johnson [10] and Golden [6] also showed that a bin-packing heuristic which uses a descending ordering performs better than other heuristics which use different orderings.

Given the list of objects and the list of boxes, the algorithm determines for each unassigned object those boxes to which it can be assigned. To determine the candidate boxes for object i, rules R1-R3 for avoiding redundant assignments are used. We also make sure that the unused capacity of each box considered is large enough for object i. Once the candidate list is set up, object i is then assigned to the lowest indexed box in the list. This assignment phase continues until all objects are assigned and a feasible solution is found.

During the course of assigning objects to boxes, some boxes may become "full." A box is full if its unused capacity is too small to accomodate any more objects. A full box and its contents can then be removed from their respective lists and a new lower bound can be computed according to (4). If, after object $i-1$ is assigned to a box, the box becomes full, then the lower bound is recomputed and denoted by $\underline{z}_i$. Of course, if the box is not full, then $\underline{z}_1 = \underline{z}_{i-1}$ for $i \geqslant 2$.

After a feasible solution is found, with the number of used boxes denoted by $\bar{z}$, the algorithm immediately backtracks to the lowest indexed object i whose $\underline{z}_i = \bar{z}$, and reassigns the object to another candidate box.

To be specific, we first define the notation used in the algorithm.

$\bar{z}$ = upper bound on the objective function value; it is initialized to ∞.

$T_j = c_j - \sum_{i=1}^{m} w_i x_{ij}$, the unused capacity of box j

U_i = Set of box indices to which object i can be assigned. U_i is determined by rules R1-R3, the condition that for every $j \in U_1$, $w_i \leqslant T_j$, and the condition that the addition of box j does not make the total number of used boxes exceed $(\bar{z} - 1)$.

$\underline{z}_i$ = lower bound after object i-1 has been assigned.

The algorithm has the following steps:

STEP 0 (Preparation): Put the boxes in descending order of their capacities (c_j), and the objects in descending order of their weights (w_i). Block boxes and object, respectively, according to their capacities and weights. Find z_1. Set $\bar{z} = \infty$ and $i = 1$.

STEP 1 (Screening): Determine U_i. If $U_i = \phi$, go to step 4.

STEP 2 (Assignment): Assign object i to the first box in U_i. Remove this box j from U_i and adjust its T_j. Compute the bound $\underline{z}_{i+1}$. If $\underline{z} \geqslant \bar{z}$, go to step 4.

STEP 3 (Forward Branching): Increment i by 1. If i does not exceed m, go to step 1. Otherwise, a feasible solution is found. Update $\bar{z}$. If $\bar{z} = \underline{z}_1$, stop because the optimal solution is found. Otherwise, find the smallest index i whose $\underline{z}_i = \bar{z}$ or which was assigned to the $\bar{z}$th box. Go to step 4.

STEP 4 (Backtracking): Decrease i by 1. If $i \geqslant 1$, go to step 2. Otherwise, the last $\bar{z}$ is the optimal value. If $\bar{z} = \infty$, the problem has no feasible solution.

4. COMPUTATIONAL EXPERIENCE

The algorithm was programmed in FORTRAN IV-G and run on a IBM 370-145. Three types of problems were randomly generated.

Type-1 problems have equal box capacities, which are equal to the maximum object weight plus 100, and the object weights are uniformly distributed between 20 and 120 in multiples of 10. Therefore, all columns in the assignment matrix are in one block and there may be as many as 10 row blocks.

Type-2 problems have equal box capacities and the object weights are uniformly distributed between 20 and 120. Therefore, there is one column block and there may be as many as 100 row blocks.

Type-3 problems have box capacities uniformly distributed between 100 and 200 in multiples of 10 and the object weights have the same distribution as Type-1 problems. Therefore there may be as many as 10 column blocks and 10 row blocks.

A total of 165 random problems of various types and various sizes were generated (see Table 1). Each problem was solved by both the algorithm and Eilon and Christofides' heuristic. Both solutions were put into subroutines and run back to back, so as to minimize the discrepancies that usually exist in computer clocking mechanisms.

TABLE 1 — *Computational Experience*

Problem		No. of Problems	Algorithm Time (in seconds)			Heuristic			
						Time (in seconds)			No. of Suboptimal Solutions
Problem Type	Size m × n		High	Low	Median	High	Low	Median	
1	30 × 30	15	0.30	0.22	0.24	0.37	0.06	0.06	2
1	40 × 40	15	1.22	0.39	0.39	0.74	0.09	0.09	2
1	50 × 50	15	0.79	0.60	0.60	1.28	0.14	0.14	2
1	60 × 60	15	1.67	0.86	0.93	2.46	0.19	0.21	2
2	30 × 30	15	450.26	0.21	0.25	0.81	0.06	0.58	3
2	40 × 40	15	0.61	0.40	0.41	1.83	0.12	1.32	4
2	50 × 50	15	46.07	0.56	0.60	2.51	0.16	1.22	3
3	30 × 30	15	0.28	0.24	0.26	0.08	0.06	0.06	0
3	40 × 40	15	0.50	0.43	0.43	0.12	0.10	0.10	0
3	50 × 50	15	0.73	0.67	0.71	0.17	0.15	0.16	0
3	60 × 60	15	1.20	1.01	1.06	0.27	0.22	0.24	0

Eilon and Christofides' heuristic [2] is an r-pass algorithm. After the boxes and the objects are numbered in the same orders as in our algorithm, it successively assigns an object to the lowest indexed box capable of accomodating it. When all the objects are assigned, one pass is complete. Then an object, which can be arbitrarily chosen, is reassigned to a box to which it was not previously assigned. The assignment of the other objects is then continued. This reassigment may be repeated up to r-1 times. In the program, when the first solution of the heuristic yields a z value greater then $\underline{z}_1$, r is set to $[\underline{z}_1/3]$, where [ ] denotes the largest integer part.

The computational experience reported in Table 1 reveals:

- The algorithm is very efficient. For the 165 test problems, only 3 required more than 1.7 s of cpu time on the IBM 370/145. For problems with 60 objects and 60 boxes, the median time was only a little over 1 s.

- Eilon and Christofides' heuristic is of high quality. The heuristic produced only 18 suboptimal solutions. For every case in which the heuristic produced a suboptimal solution, the associated value exceeded the optimal by only 1.

- The lower bound computed by (4) is generally effective. One measure of the tightness of bounds is the ratio of the bound to the optimal value [4]. For the 165 problems, the lowest ratio ($\underline{z}_1/\bar{z}$) is 81.5% and the average ratio is 98.5%.

- The performance of the algorithm is not greatly affected by the bound. For example, in the three problems that took more than 1.7s for the algorithm, the bounds were all one less than the respective optimal values.

5. DISCUSSION

Further experiments were conducted to improve the efficiency of the algorithm. Specifically, in light of the recent bounding techniques for integer programming [4], we wanted to see if bounds better (higher) than those found by (4) could be obtained.

Following the Lagrangean relaxation approach suggested by Geoffrion [4], we constructed a Lagrangean dual formulation of (p). The obtained Lagrangean dual problem is as follows:

(L) $$\operatorname*{Max}_{\lambda} \operatorname*{Min}_{x,j} \sum_{j=1}^{n} y_j - \sum_{i=1}^{m} \lambda_i \left(\sum_{j=1}^{n} x_{ij} - 1 \right)$$

subject to

(2) $$\sum_{i=1}^{m} w_i x_{ij} \leqslant c_j y_j$$

(3) $$x_{ij} = 0 \text{ or } 1,\ y_j = 0 \text{ or } 1 \text{ for all } i,\ j.$$

Given a value for each λ_i, (L) decomposes into a collection of 0-1 knapsack problems, one for each box j. That is, for each j, we solve

(L) $$(L)_j \qquad \operatorname{Max} \sum_{i=1}^{m} \lambda_i x_{ij}$$

subject to

(5) $$\sum_{i=1}^{m} w_i x_{ij} \leqslant c_j$$

$$x_{ij} = 0 \text{ or } 1 \text{ for all } i,\ j.$$

Let x_{ij}^* be the optimal solution and $\gamma_j = \sum_{i=1}^{m} \lambda_i x_{ij}^*$ be the optimal value of $(L)_j$. Then for each j, the optimal decisions will be

$$y_j = 1 \text{ and } x_{ij} = x_{ij}^* \text{ for } i = 1, \ldots m, \text{ if } \gamma_j > 1, \text{ and}$$

(6)
$$y_j = 0 \text{ and } x_{ij} = 0 \text{ for all } i, \text{ otherwise.}$$

The explanation for (6) is that if $\gamma_j > 1$, then by letting $y_j = 1$ (use box j) and assigning objects whose $x_{ij}^* = 1$, to box j, we can reduce the objective function value of (L).

To find the optimal λ (that which maximizes (L)), we use the subgradient ascent method [8]. At iteration $t \geqslant 0$, define the subgradient vector (S_i^t) to be

$$S_i^t = \sum_{j=1}^{n} x_{ij}^* - 1.$$

x_{ij}^* is the optimal solution of $(L)_j$ based on (λ_i^t). Then the next λ vector is, for each component,

(7)
$$\lambda_i^{t+1} = \lambda_i^t + \theta S_i^t,$$

where θ is a step size determined in the same manner described in [8]. Once λ^{t+1} is computed, we go back to solve knapsack problems $(L)_j$.

Five problems of 30 objects and 30 boxes were tested. For each of the five problems, the above procedure did not produce a bound better than that found by (4) after 300 iterations and nearly 1 cpu minute. This experience suggests that for the test problems considered in this paper, the bound derived from (4) may be the best, in terms of quality and ease of computation.

REFERENCES

[1] Brown, J. R., "Subductive Programming and Chromatic Scheduling," Ph. D. dissertation, Massachusetts Institute of Technology (1970).

[2] Eilon, S., and N. Christofides, "The Loading Problem," Management Science *17*, 259-268 (1971).

[3] Eilon, S., C. D. T. Watson-Gandy, and N. Christofides, *Distribution Management*, (Hafner, N.Y., 1971).

[4] Geoffrion, A., "Lagrangean Relaxation for Integer Programming," in *Mathematical Programming Study 2, Approaches to Integer Programming* M. L. Balinski, Ed. 82-114 (North-Holland Publishing Co., Amsterdam 1974).

[5] Gilmore, P., and R. Gomory, "Multistage Cutting Stock Problems of Two and More Dimensions," Operations Research, *13* 94-120 (1965).

[6] Golden, B. L., "Approaches to the Cutting Stock Problem," *AIIE Transactions, 2*, 265-274 (June 1976).

[7] Greenberg, I., "Application of the Loading Algorithm to Balance Workloads," *AIIE Transactions, 4*, 337-339 (1972).

[8] Held. M., P. Wolfe, and H. Crowder, "Validation of Subgradient Optimization" Mathematical Programming *6*, 62-88 (1974).

[9] Ingargiola, G., and J. F. Korsh, "An Algorithm for the solution of 0-1 Loading Problems," Operations Research *23*, 1110-1119 (1975).

[10] Johnson, D., "Fast Algorithms for Bin Packing," Journal of Computer and System Science, *8*, 272-314 (1974).

[11] Ross, G. T., and R. M. Soland, "A Branch and Bound Algorithm for the Generalized Assignment Problem," Mathematical Programming *8*, 91-103 (1975).

A HEURISTIC NETWORK PROCEDURE FOR THE ASSEMBLY LINE BALANCING PROBLEM

Peter A. Pinto

Bowling Green State University,
Bowling Green, Ohio

David G. Dannenbring and Basheer M. Khumawala

University of North Carolina
Chapel Hill, North Carolina

ABSTRACT

Proposed is a Heuristic Network (HN) Procedure for balancing assembly lines. The procedure uses simple heuristic rules to generate a network which is then traversed using a shortest route algorithm to obtain a heuristic solution. The advantages of the HN Procedure are: a) it generally yields better solutions than those obtained by application of the heuristics, and b) sensitivity analysis with different values of cycle time is possible without having to regenerate the network. The rationale for its effectiveness and its application to problems with paralleling are presented. Computational experience with the procedure on up to 50 task test problems is provided.

1. INTRODUCTION

Assembly line production methods are extensively used in high-volume manufacturing. Because of the typically high level of output in line production, even a small reduction in per-unit cost results in substantial overall savings. Since Salveson [13] first formulated the problem in 1956, several researchers [1,2,4,5] have suggested both exact and heuristic procedures to the well-known single-model, assembly line balancing (SMALB) problem.

More recently, extensions to the SMALB problem, such as mixed model lines [8,14], stochastic balancing [7], and paralleling [2,3,11,12] of assembly lines, have been proposed. The mixed assembly line has several variations of the same general product intermixed on one line. Stochasticity allows for variability in the task times and considers the consequences of incompletion of tasks at stations. The paralleling extension, which is of special interest in this paper, allows tasks to be performed at more than one station.

It is important to note that the solution procedures suggested for the aforementioned extensions to the SMALB problem normally require repeated solution of many SMALB subproblems. Thus, their success relies heavily on being able to solve the basic SMALB problem efficiently. This paper presents such an efficient procedure.

The remainder of the paper is organized as follows. The exact solution procedure to the SMALB problem proposed by Gutjhar and Nemhauser [4] is reviewed in the next section. In Section 3 the Heuristic Network (HN) Procedure, which we have derived from the Gutjhar and Nemhauser algorithm, is described. The basis for the effectiveness of the HN Procedure is also examined. The procedure is then applied to an example problem. Section 4 presents computational experience with the HN Procedure and suggests directions for further study.

2. MATHEMATICAL FORMULATION OF THE SMALB PROBLEM

Let $cr = T/n$ be the production constrained cycle time where n is the number of units desired to assemble in time T,

K = set of tasks to be performed on any one unit of the product

$= \{1, 2, \ldots k\}$

$t(i)$ = the time required to perform task $i \in K$

Q = set of stations in the assembly line

$= \{1, 2, \ldots q\}$

$A(j)$ = set of tasks contained in the jth station.

$W(j)$ = total work content in the jth station; namely, sum of the task times of all tasks contained in the jth station.

Then, the objective of the SMALB problem is to

(2.1) Minimize Idle Time $= Z = q^*cr - \sum_{i \in K} t(i)$

subject to:

(a) $\bigcup_j A(j) = K$

(All tasks have to be performed.)

(b) $A(j) \cap A(i) = \phi$ (the empty set)

(No task can be assigned more than once.)

(c) $W(j) = \sum_{k \in A(j)} t(k) \leqslant cr; j \in Q$

(The work content in any station cannot exceed the cycle time.)

(d) If $x < y$ and $x \in A(i)$, $y \in A(j)$; then $i \leqslant j$

(If task x precedes task y, then y cannot be allocated to a station that precedes the one to which x is assigned.)

It is well known that the objective of the SMALB problem given above is often reduced to one of two alternate forms:

(2.2) Minimize $Z = q$, given cr, or

(2.3) Minimize $Z = cr$, given q.

This is because $\sum_{i \in K} t(i)$ (the total work content) is a given constant in expression (2.1), leaving q or cr as the only variables to be minimized. It the production constrained cycle time, cr, is known, we can attempt to minimize q as in expression (2.2). Similarly, given q, we have cr to be minimized as in expression (2.3). Henceforth, when we refer to the SMALB problem we will be referring to minimizing expression (2.2).

The Gutjhar and Nemhauser algorithm [4] begins by generating a directed network based on the precedence diagram for the SMALB problem. The network is constructed with nodes representing a collection or subset of tasks that can be performed in some order without prior completion of any task not in the subset. Let $C(i)$ and $t[C(i)]$ represent the tasks at node i and the sum of their times, respectively. Clearly, the source and sink nodes are given by $C(0) = \phi$, $t[C(0)] = 0$, $C(r) = K$ (the complete set of tasks), and $t[C(r)] = \sum_{j \in K} t_j$.

In the network a directed arc (ij) is defined if and only if $C(i) \epsilon C(j)$ and $t[C(j)] - t[C(i)] \leqslant cr$. Clearly, the arc corresponds to a station assignment of all tasks contained in $C(j)$ but not in $C(i)$.

As a result of this definition of the network, there is a one-to-one correspondence between paths from the source node (node 0) to the sink node (node r) and feasible assignments of tasks to stations. Thus, a path containing q arcs required to traverse the network from node 0 to node r is equivalent to q work stations each with a cycle time of cr. Consequently, idle time equals

$$qcr - \sum_{i \in K} t(i).$$

Thus, finding a solution to the SMALB problem is equivalent to finding the shortest path of this network.

The major difficulty experienced with this algorithm is that the number of nodes generated increases exponentially with problem size [4], which leads to heavy computational requirements of both time and storage.

3. THE HEURISTIC NETWORK

3.1 Procedure

While generating all possible nodes guarantees an optimal solution, a heuristic solution can be obtained by generating a limited number of nodes. The quality of the solution can be enhanced by using heuristic rules to generate relevant and promising nodes. In this research we use the following four popular heuristics [1,6,9,15]: (1) Ranked Positional Weight, (2) Largest Task Time, (3) Smallest Task Time, and (4) Random Assignment. Unlike the first three rules, the Random Assignment rule can be applied a number of times to generate additional nodes.

Each heuristic can be used to generate a solution which is equivalent to a set of nodes in the directed network. After each heuristic is applied independently, the sets of nodes can be

combined to form a composite network. This allows consideration of additional arcs and as a result the possibility of a shorter path and, hence, a better solution.

When the first heuristic rule is applied, a total $k + 1$ nodes are obtained for a k task problem in the composite network. Next, by applying the second heuristic rule we obtain an additional $k - 1$ nodes, as the sink and source nodes are clearly common for all heuristic rules. Continuing this process, if n heuristic rules are applied we develop a total of $[(n - 1)\ (k - 1) + (k + 1] = n(k - 1) + 2$ nodes in the composite network. Of course, the number of nodes in the composite network can be further reduced, by eliminating duplicate nodes during the generation process.

Additional arcs are then added between any two nodes i and j for which $C(i) \in C(j)$ and $t[C(j)] - t[C(i)] \leqslant cr$.

For example, in Fig. 1 two heuristically generated networks are shown. The shortest path (equivalent to stations of length less than or equal to the required cycle time of 50) consists of 4 arcs — 0, 1, 2, 4, 5 or 0, 1, 2, 3, 5 and 0, 6, 7, 8, 9, respectively. Consider a composite network obtained by merging the two networks. As shown in Fig. 2, the augmented network contains 2 additional arcs: between nodes 1 and 8 and nodes 8 and 4. Note that the shortest paths in the composite network consist of only 3 arcs (stations): 0, 1, 8, 5 (and 0, 1, 8, 9). This demonstrates that the composite network may in fact yield a better solution than when the heuristics are used separately.

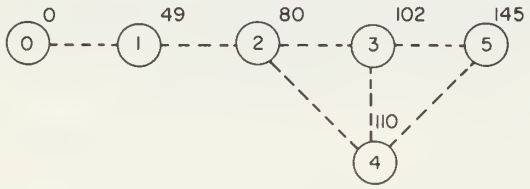
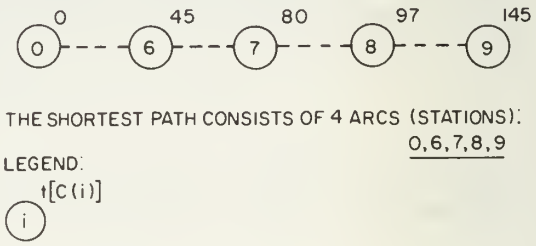


FIGURE 1(a). Network generated by using a heuristic.



(b). Network generated by using another heuristic.

FIG. 1(a) AND 1(b), AUGMENTED BY ADDITIONAL ARCS.

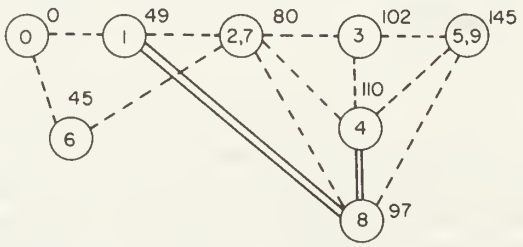


FIGURE 2. Composite network.

3.2 An Illustrative Example

Consider the 11-task SMALB problem shown in Fig. 3. Let the cycle time, *cr*, be 50. For purposes of illustration, only three heuristic rules have been used to obtain the composite network. Specifically, the Random Assignment rule was applied twice, and the Largest Task Time rule employed once. Table 1 shows the three solutions generated by the application of the three heuristic rules. The work content times for the stations are also indicated. Table 2 details how the nodes were generated for one of the heuristics.

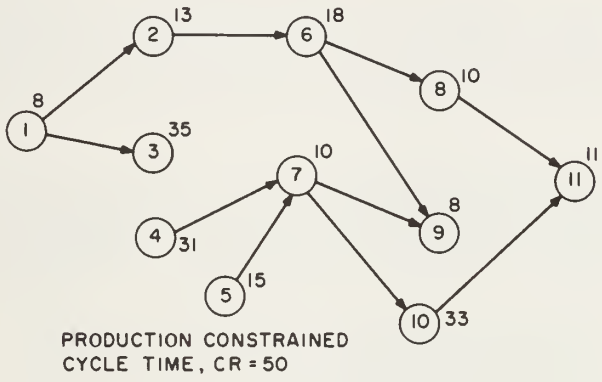


FIGURE 3. Illustrative problem.

Table 1. — *Generation of Heuristic Solutions*

Solution No.	Heuristic Rule Used	Task Assignments
1	Random $C(i)$	1, 2, 6, 8 | 4,5 | 3,7 | 9,10 | 11
	$t[C(i)]$	---→49 | 46 | 45 | 41 | 11
2	Largest $C(i)$	1, 3 | 4, 5 | 2, 6, 7, 9 | 10, 8 | 11
	Task Time $t[C(i)]$	43 | 46 | 49 | 43 | 11
3	Random $C(i)$	1, 3 | 5, 2, 6 | 8, 4 | 7, 9 | 10, 11
	$t[C(i)]$	43 | 46 | 41 | 18 | 44

Legend:
---→ : station assignments corresponding to the shortest path through the composite network

Note that each heuristic will yield two common nodes: the source and the sink nodes. For the three heuristic solutions we have 32 nodes labelled from 0 to 31, nodes 0 and 11 being the common source and sink nodes, respectively. The composite network is illustrated in Fig. 4.

Each of the three heuristics considered separately yields a five-station solution. However, the composite network shows that a four-station solution is possible, as shown in Table 1 and Fig. 4.

When the composite network is traversed with arc length equal to the desired cycle time of 50, a minimum of four arcs are required. As shown in Fig. 4, the first station is equivalent to arc (0,4), the second to arc (4,27), the third to arc (27,29), and the fourth to arc (29,11). The assignment of tasks to stations is then obtained by "backtracking," i.e., tasks are assigned to

Table 2. — *Tasks in Nodes for Sequence No. 1, Using Random Assignment Heuristic*

Node No. (i)	Tasks	Work content t[C(i)] for Node
0	0	0
1	1	8
2	1,2	21
3	1,2,6	39
4	1,2,6,8	49
5	1,2,6,8,4	80
6	1,2,6,8,4,5	95
7	1,2,6,8,4,5,3	130
8	1,2,6,8,4,5,3,7	140
9	1,2,6,8,4,5,3,7,9	148
10	1,2,6,8,4,5,3,7,9,10	181
11	1,2,6,8,4,5,3,7,9,10,11	192

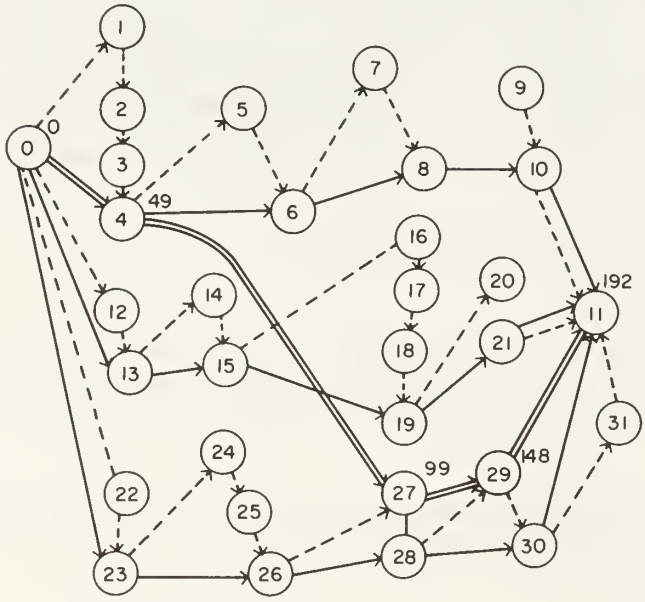


FIGURE 4. Composite network.

TASKS IN STATION 4 = C(11) − C(29) = {10, 11}

TASKS IN STATION 3 = C(29) − C(27) = {4,7,9}

TASKS IN STATION 2 = C(27) − C(4) = {3,5}

TASKS IN STATION 1 = C(4) − C(0) = {1,2,6,8}

the last station first. To illustrate, consider the assignment of tasks to station 2. As shown in Table 1, node 27 corresponds to the set of tasks $C(27) = (1,3,5,2,6,8)$ with work content of 99, while node 4 corresponds to $C(4) = (1,2,6,8)$ with work content of 49. The set of tasks assigned to station 2 are those that appear in set $C(27)$ but not in $C(4)$. This yields tasks 3 and 5 with a station work content of 50. The assignment of tasks to stations which achieves the four-station balance is shown in Fig. 4.

4. COMPUTATIONAL EXPERIENCE WITH THE HN PROCEDURE

The HN Procedure was programmed in Fortran IV and was tested on problems of varying difficulty that have appeared in the literature. The program and full details of the problem sets can be found in Pinto [12]. The largest problem solved was a 50-task problem due to Mansoor [10].

Table 3 provides the computational results. Balance efficiencies were above 96% for all problems but one. Computation times, of course, varied with the parameters of the problem at hand. The computational efficiency is improved when the HN Procedure is applied to multiple iteration problems, as described in the next section.

Table 3. — *Computational Results for a Set of SMALB Problems Using the HN Procedure*

No. of Tasks	No. of Stations	Minimum Cycle Time Possible	Cycle Time Achieved	Balance Efficiency	CPU Time in Seconds (IBM 360/75)
11*	4	50	50	96.5%	0.22
15**	5	—	39	90%	0.45
17 (Thomopolous)	3	126	126	100%	0.90
45 (Kilbridge)	3	184	184	100%	5.2
29 (Buxey)	7	48	50	97%	1.1
50 (Mansoor)	5	—	472	98%	3.2
50 (Mansoor)	10	—	244	96.8%	9.4

*See Fig. 3

**—randomly generated problem

4.1 Computational Experience With HN Procedure for Multiple Interation Problems

Multiple iteration line balancing problems, such as paralleling problems [2,11,12], are those in which several SMALB subproblems are required to be solved. Since the number of these subproblems often increases exponentially, computational effort in solving individual subproblems becomes very important.

It is convenient to consider the HN Procedure as consisting of two stages. In Stage 1 the nodes of the network are generated, while in Stage 2 the arcs are added and the network traversed to obtain task to station assignments. When applied to multiple iteration problems,

significant computational savings are achieved because Stage 1 is required to be performed only once.

To illustrate, consider the branch and bound method of solving the paralleling of stations problem [12]. Each successive iteration of the branch and bound technique requires the solution of an SMALB problem. The difference between two consecutive subproblems is that each successor subproblem will have one additional station paralleled.

This is conveniently handled by the HN Procedure, since the paralleling of a station only affects Stage 2 computations. Table 4 shows the application of the HN procedure to the paralleling of stations problem, demonstrating that significant computational savings are possible.

Table 4. — *Computation Time for Paralleling of Station Problems*

No. of Tasks	Iteration	No. of Stats.	Cycle Time Achieved	Time to Generate Network (Stage 1)	CPU Time in Seconds (IBM 360/75) Time to find Shortest Route (Stage 2)	Cumulative Time
11 (Fig. 3)	1	4	52	0.15	0.08	0.23
	2	3	50	—	0.09	0.32
	3	3	50	—	0.10	0.42
29 (Buxey)	1	7	50	1.3	1.1	2.4
	2	6	48	—	1.2	3.6
	3	6	48	—	1.4	5.0

REFERENCES

[1] Arcus, A. L., "COMSOAL: A Computer Method of Sequencing Operations for Assembly Lines," Parts I and II, in *Readings in Production and Operations Management*, E. S. Buffa, ed. John Wiley and Sons, New York (1966).

[2] Buxey, G. M., "Assembly Line Balancing with Multiple Stations," Management Science, *20*, (6) (February, 1974).

[3] Freeman, J. R., and J. V. Jucker, "The Line Balancing Problem," Journal of Industrial Engineering, *18*, (6) (June, 1967).

[4] Gutjhar, A. L., and G. L. Nemhauser, "An Algorithm for the Line Balancing Problem," Management Science *11*, (2) (November, 1964).

[5] Held, M., R. M, Karp, and R. Sharesian, "Assembly Line Balancing-Dynamic Programming with Precedence Constraints," Operations Research, *11*, (3) (May-June, 1963).

[6] Ignall, E. J., "A Review of Assembly Line Balancing," Journal of Industrial Engineering, *16*, (4) (July-August, 1965).

[7] Kottas, J. F., and H. S. Lau, "A Cost-Oriented Approach to Stochastic Line Balancing," AIIE Transactions *9*, (2) (March-April, 1974).

[8] Macaskill, J. L. C., "Production-Line Balances for Mixed-Model Lines," Management Science, *19*, (4) (December, 1972).

[9] Mansoor, E. M., "Assembly Line Balancing — An Improvement on the Ranked Positional Weight Technique," Journal of Industrial Engineering, *15*, (2) (March-April, 1964).

[10] Mansoor, E. M., "Solving Large Single-Model Assembly Lines," AIIE Transactions *9*, (2) (March-April, 1974).

[11] Pinto, P. A., D. G. Dannenbring, and B. M. Khumawala, "A Branch and Bound Algorithm for Assembly Line Balancing with Paralleling," International Journal of Production Research *13*, (2) (1975).

[12] Pinto, P. A., "Assembly Line Balancing With Paralleling," Unpublished Ph.D. thesis, University of North Carolina, Chapel Hill, (1975).

[13] Salveson, M. E., "The Assembly Line Problem," Journal of Industrial Engineering *6*, (3) (May-June, 1955).

[14] Thomopoulos, N. T., "Mixed-Model Line Balancing with Smoothed Station Assignments," Management Science *16*, (9) (May, 1970).

[15] Tonge, F. M., "Summary of a Heuristic Line Balancing Procedure," Management Science *7*, (1) (October, 1960).

PSEUDO-MONOTONIC INTERVAL PROGRAMMING

C. R. Bector and S. K. Bhatt*

Faculty of Administrative Studies
University of Manitoba, Canada

ABSTRACT

A pseudo-monotonic interval program is a problem of maximizing $f(x)$ subject to $x \in X = \{x \in R^n \mid a \leqslant Ax \leqslant b;\ a,\ b \in R^m\}$ where f is a pseudo-monotonic function on X, the set defined by the linear interval constraints. In this paper, an algorithm to solve the above program is proposed. The algorithm is based on solving a finite number of linear interval programs whose solutions techniques are well known. These optimal solutions then yield an optimal solution of the proposed pseudo-monotonic interval program.

INTRODUCTION

Since the publication of the first paper on the method of solving a linear interval program [4], there have been several attempts to look into possible generalizations. A linear fractional interval program is an obvious generalization, as it is rooted in Charnes and Cooper's observation that a linear fractional program is equivalent to a linear program [7]. Subsequently, Charnes and Cooper [8], Buhler [6], and Bector and Bhatt [1] gave different methods for linear fractional interval programming, the last two mainly reducing the linear fractional interval program to an equivalent linear interval program.

Further generalization of the problem of maximizing the nonlinear interval programs does not seem obvious due to difficulty in working out suitable sufficient optimality criteria. This paper is but a step in this direction.

DEFINITIONS [9]. Let $f:D \to R$ be a differentiable function on D, an open subset of R^n. Let $S \subset D$ and $x^u,\ x^v \in S$. Then on S, f is said to be

(i) pseudo-convex if
$\nabla f(x^v)'(x^u - x^v) \geqq 0 \to f(x^u) \geqq f(x^v)$

(ii) quasi-convex if
$f(x^u) \leqq f(x^v) \to \nabla f(x^v)'(x^u - x^v) \leqq 0$

(iii) pseudo-concave if $-f$ is pseudo-convex on S.

(iv) pseudo-monotonic if f is both pseudo-convex and pseudo-concave on S.

*On leave from Indian Statistical Institute, Calcutta, India.

The pseudo-monotonic interval program under consideration is:

(P) maximize $f(x)$
subject to $x \in X = \{x \in R^n \mid a \leqq Ax \leqq b;\ a,\ b \in R^m\}$

where A is an $m \times n$ matrix and f is pseudo-monotonic on X. Also it is assumed that

A1. X is bounded.
A2. A is of full row rank.

THE ALGORITHM (LINEARIZATION TECHNIQUE)

The proposed algorithm to solve the pseudo-monotonic interval program (P) is based on the following theorem.

THEOREM 1 (Kortanek and Evans [9]): Let f be a pseudo-concave function on a closed, convex set C and $x^* \in C$. Consider the two programs:

I: maximize $f(x)$, subject to $x \in C$.
II: maximize $\nabla f(x^*)'x$, subject to $x \in C$.

Then x^* is an optimal solution of program I if and only if X^* is an optimal solution of program II.

As the objective function of the program (P) is pseudo-monotonic and hence pseudo-concave, the algorithm aims at finding an $x^* \in X$, the feasible set of (P), which satisfies Theorem 1. This is done by successively generating extreme points of the polyhedral constraint set X and at each step using the stopping rule given at the end of the description of the algorithm below. The stopping rule is based on Theorem 1.

To start the algorithm, an initial extreme point of the constraint set is required. For this, choose an $x^0 \in R^n$ such that $\nabla f(x^0) \neq 0$. Now consider the following linear interval program:

(L_0) Maximize $\nabla f(x^0)'x$
subject to $x \in X$.

It would be interesting to note that an optimal solution of (L_0) exists, as the object function is continuous and the constraint set is compact by assumption A1. Then by Lemma 6 of Ref. [4], $f(x^0) \in N(A)^{\perp}$, the orthogonal complement of the null space of A. This, coupled with assumption A_2, makes (L_0) satisfy the requirements to use the method of Ben-Israel and Charnes [4] to solve (L_0). A solution of (L_0) then gives an extreme point solution x^1 of (L_0) (see example 2 at the end of the paper).

The algorithm can now be described through the following steps.

STEP 1: Let x^1 be an initial extreme point of X. Set $i = 1$. Solve by Ben-Israel and Charnes' method the linear interval program:

$$(L_i) \qquad \text{maximize} \qquad \nabla f(x^i)'x, \text{ subject to } x \in X.$$

Let x^{i+1} be an extreme point optimal solution of (L_i). Note: The remark made for the program (L_0) above is also valid for each (L_i) in the sense that $\nabla f(x^i) \in N(A)^{\perp}$, satisfying the assumptions of Ben-Israel and Charnes' method of solving linear interval programs [4].

STEP 2 : (i) If $x^{i+1} \in \{x^1, x^2, \ldots, x^i\}$, stop. Suppose $x^{i+1} = x^j$, for some j, $1 \leqq j \leqq i$. Then $x^j, x^{j+1}, \cdots, x^i$ solve the program (P) due to theorem 2 below.

(ii) If $x^{i+1} \notin \{x^1, x^2, \ldots, x^i\}$, go to step 1 with $i = i + 1$.

THEOREM 2. (The stopping rule): If in step 2 of the algorithm, $x^{i+1} \in \{x^1, x^2, \ldots, x^i\}$, i.e. $x^{i+1} = x^j$ for some j, $1 \leqq j \leqq i$, then $x^j, x_{j+1}, \ldots, x^i$ solve the program (P).

PROOF. Without loss of generality, let x^{k+1} solve (L_k), $1 \leqq k \leqq i$. That is, $\nabla f(x^k)' x^{k+1} \geqq \nabla f(x^k)' x$ for all x in X. In particular, $\nabla f(x^k)' x^{k+1} \geqq \nabla f(x^k)' x^k$ as $x^k \in X$. That is, $\nabla f(x^k)'(x^{k+1} - x^k) \leqq 0$, implying that $f(x^{k+1}) \geqq f(x^k)$ as f is a pseudo-convex function. Hence

$$f(x^1) \leqq f(x^2) \leqq \cdots \leqq f(x^i) \leqq f(x^{i+1}). \tag{1}$$

Now it is given that $x^j = x^{i+1}$ maximizes (L_j), for some j, $1 \leqq j \leqq i$. If $j = i$, i.e., $x^j = x^i$, then by theorem 1, x^i solves (P). If $1 \leqq j < i$, then

$$\nabla f(x^j)' x^j \geqq \nabla f(x^j)' x \text{ for all } x \in X.$$

But $x^i \in X$, so that

$$\nabla f(x^j)'(x^i - x^j) \geqq 0, \text{ implying that} \tag{2}$$

$$f(x^i) \geqq f(x^j) \text{ as } f \text{ is pseudo-convex.} \tag{3}$$

Combine (1) and (3) to get

$$f(x^j) = f(x^{j+1}) = \cdots = f(x^i). \tag{4}$$

Since f is pseudo-convex, it is quasi-convex also [10]; therefore, $f(x^j) = f(x^i) \rightarrow \nabla f(x^j)'(x^i - x^j) \leqq 0$, or

$$\nabla f(x^j)' x^j \leqq \nabla f(x^j)' x^i. \tag{5}$$

From (2) and (5) and the fact that $x^j = x^{i+1}$,

$$\nabla f(x^j)' x^j = \nabla f(x^j)' x^{i+1}, \tag{6}$$

which implies that x^j maximizes (L_j). Then by theorem 1, x^j also maximizes (P) and so do $x^j, x^{j+1}, \cdots x^{i-1}$ due to (4) above.

CONVERGENCE

The algorithm generates a sequence of extreme points of the convex constraint set X until one of the extreme points generated earlier is repeated. At this state, optimality is reached due to theorem 2. Since X is bounded, it has finitely many extreme points and hence convergence occurs in finitely many iterations. This observation is in accordance with the result of Bela Martos [2] applied to this situation that a maximum of a pseudo-monotonic function lies at one of the extreme points of the compact convex constraint set X. In fact, the result of Bela Martos is for differentiable quasi-monotonic functions, but then pseudo-monotonicity → quasi-monotonicity for differentiable functions.

REMARKS

This paper is an extension of the linearization technique used to solve a pseudo-monotonic program [5] and an interval linear fractional program in a finite number of iterations

[1]. In the case of a pseudo-monotonic program which is maximizing a pseudo-monotonic function subject to the linear constraints $Ax \leqq b$, $x \geqq 0$, the programs (L_t) are the ordinary linear programs which can be solved by simplex method.

ILLUSTRATION

EXAMPLE 1 The interval linear fractional program: maximize $(c'x + \alpha)/(d'x + \beta)$, subject to $x \in X = \{x \in R^n | a \leqq Ax \leqq b,\ a,\ b \in R^m\}$ and $d'x + \beta > 0$ for $x \in X$, is an example of a pseudo-monotonic interval program as a linear fractional function. $(c'x + \alpha)/(d'x + \beta)$ on X is also a pseudo-monotonic function on X [3].

EXAMPLE 2†: maximize

$$f(x) = \frac{x_1 + 2 - x_2\sqrt{(x_1 + 2)^2 + (x_2)^2 - 1}}{(x_1 + 2)^2 + (x_2)^2},$$

subject to

$$\begin{aligned} -1 &\leqq x_1 & &+ x_3 \leqq 2 \\ 1 &\leqq & x_2 & \leqq 5. \\ 0 &\leqq & & x_3 \leqq 8 \end{aligned}$$

Here

$$a = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix},\ b = \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix},\ A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ \text{and } x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

$N(A)$ is spanned by zero vector in R^3 [4]. Denote by X^0 the set of solutions of the interval constraints. Now for any x,

$$\nabla f(x) = \begin{bmatrix} -\dfrac{(x_1 + 2 - x_2\xi)[x_2 + (x_1 + 2)\xi]}{(\xi^2 + 1)^2\xi} \\ -\dfrac{[x_2 + (x_1 + 2)\xi]^2}{(\xi^2 + 1)^2\xi} \\ 0 \end{bmatrix}$$

where $\xi = \sqrt{(x_1 + 2)^2 + (x_2)^2 - 1}$. Denote

$$\begin{bmatrix} -(x_1 + 2 - x_2\xi)[x_2 + (x_1 + 2)\xi] \\ -[x_2 + (x_1 + 2)\xi]^2 \\ 0 \end{bmatrix} \text{ by } \overline{\nabla} f(x).$$

Then $\overline{\nabla} f(x)$ is a positive multiple of $\nabla f(x)$, and therefore maximizing $\nabla f(x')'x$ for any $x' \in X^0$ over all $x \in X^0$ is equivalent to maximizing $\overline{\nabla} f(x')'x$ over $x \in X^0$ in the sense that both the maximizing problems admit same optimal solutions. It is easy to see that $x^{-1} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ is an extreme point of X^0. Also, $\overline{\nabla} f(x)' = \begin{bmatrix} 0 \\ -4 \\ 0 \end{bmatrix}$.

†The objective function of example 2 is due to Bela Martos [3].

STEP 1: Solve the linear interval program

$$(L_1) \quad \begin{array}{ll} \text{maximize} & \overline{\nabla} f(x^1)'x \equiv -4x_2 \\ \text{subject to} & x \in X^0. \end{array}$$

The method of Ben-Israel and Charnes gives

$$\begin{bmatrix} -\lambda + 2(1-\lambda) - 8\mu \\ 1 \\ 8\mu \end{bmatrix}, \quad 0 \leqq \lambda \leqq 1$$

and $0 \leqq \mu \leqq 1$ as the optimal solutions of (L_1). The extreme point optimal solutions in particular can be obtained by taking λ and μ as 0 or 1 which are

$$\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -6 \\ 1 \\ 8 \end{bmatrix} \text{ and } \begin{bmatrix} -9 \\ 1 \\ 8 \end{bmatrix}.$$

Let $x^2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ be taken as an extreme point optimal solution of (L_1).

STEP 2: $x^2 \notin \{x^1\}$. So go to step 1 and solve the linear interval program:

$$(L_2) \quad \begin{array}{ll} \text{maximize} & \overline{\nabla} f(x^2)'x \equiv -289x_2 \\ \text{subject to} & x \in X^0. \end{array}$$

(L_2) has same optimal solutions as (L_1). Therefore $x^3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ is an extreme point optimal solution of (L_2).

STEP 3: $x^3 \in \{x^1, x^2\}$ as $x^3 = x^2$. Hence $x^2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ gives the optimal value of $f(x)$ which is 0.

ACKNOWLEDGMENT

Thanks are due to the referee for helpful suggestions.

REFERENCES

[1] Bector, C. R., and Bhatt, S. K., "A Linearization Technique for Solving Interval Linear Fractional Programs," Presented at the 5th Manitoba Conference on Numerical Mathematics (October 1975).

[2] Bela Martos, "The Direct Power of Adjacent Vertex Programming Methods," Management Science *12*(3), 241-252, (1965).

[3] Bela Martos, *Nonlinear Programming — Theory and Methods,* (North-Holland Publ. Co. 1975).

[4] Ben-Israel and Charnes, "An Explicit Solution of a Special Class of Linear Programming Problems," Operations Research *16* (6), 1166-1175 (1968).

[5] Bhatt, S. K., "Generalized Pseudo-Convex Programming and Optimal Control," Thesis submitted to the Department of Mathematics, I.I.T. Kanpur, India (1971).

[6] Buhler, W., "A Note on Fractional Interval Programming," Zeitschrift fur Operations Research, *19* (1), 29-36 (1975).

[7] Charnes, A., and Cooper, W. W., "Programming With Linear Fractional Functionals," Naval Research Logistics Quarterly, *9* (3-4), 181-186 (1962).

[8] Charnes, A., and Cooper, W. W., "An Explicit General Solution in Linear Fractional Programming," Naval Research Logistics Quarterly, *20* (3), 449-467 (1973).

[9] Kortanek, K. O., and Evans, J. P., "Pseudo-Concave Programming and Lagrange Regularity," Operations Research *15* (5), 882-891 (1967).

[10] Mangasarian, O. L., "Pseudo-Convex Functions," SIAM Journal on Control, Series B, (1965).

A TARGETING MODEL THAT MINIMIZES COLLATERAL DAMAGE

Jeffrey H. Grotte

Institute for Defense Analyses
Arlington, Virginia

ABSTRACT

This paper considers the problem of allocating weapons to achieve targeting objectives while simultaneously minimizing aggregate damage to surrounding nonmilitary facilities, each of which has an upper limit to the damage it is permitted to incur. A model is formulated which assumes only that damage to individual targets or associated facilities does not decrease as the number of allocated weapons increases. An implicit enumeration algorithm, based on that of Lawler and Bell is described that yields optimal integer solutions. An example is presented.

1. INTRODUCTION

One of the assumptions behind the argument to employ *counterforce* targeting of strategic weapons (the targeting of an enemy's strategic capability), as opposed to *countervalue* targeting (the objective of which is the destruction of population and economy), is that sufficient destruction of strategic targets can be achieved without causing appreciable damage to the surrounding nonstrategic facilities. This paper presents a model which addresses the following two questions: Given a collection of weapons, potential aimpoints, and a configuration of strategic targets — each being assigned a *minimum* level of damage, and nonstrategic facilities — each having a *maximum* level of permissible damage.

(A) Is there an assignment of weapons to aimpoints (an *allocation*) that satisfies the above two sets of constraints?

(B) Of all allocations satisfying the above two sets of constraints, which is the one (or *a* one) that minimizes the (perhaps weighted) sum of the damage to the nonstrategic facilities?

2. MATHEMATICAL FORMULATION

The fundamental elements of the model are M strategic targets, henceforth called simply "targets," N nonstrategic facilities, or "nontargets," I different weapon types, and J potential aimpoints to which any weapon can be directed. An allocation $\underset{\sim}{z}$ is the matrix $\{z_{i,j} | i = 1, ..., I; j = 1, ..., J\}$ where $z_{i,j}$, an integer, is the number of weapons of type i allocated to aimpoint j.

For each target m, we suppose a real-valued response function $f_m(\underset{\sim}{z})$ which represents the damage to target m from allocation $\underset{\sim}{z}$. We require that $f_m(\underset{\sim}{z})$ be monotonically nondecreasing

in each component of $\underset{\sim}{z}$, which is an implicit assumption that, given any allocation, the allocation of additional weapons does not result in less damage to any target. Each target m is assigned a real number c_m which is the minimum damage requirement (targeting objective); i.e., for an allocation $\underset{\sim}{z}$ to be feasible, it must satisfy $f_m(\underset{\sim}{z}) \geqslant c_m$, $m = 1, \ldots, M$.

Similarly, for each nontarget n there is a response function $g_n(\underset{\sim}{z})$, monotonically nondecreasing in each component of $\underset{\sim}{z}$, and a real number d_n denoting the maximum damage permitted to this nontarget. Further, each nontarget n is assigned a nonnegative weight, or value, λ_n.

The nonnegative integer w_i is the number of weapons of type i available for allocation.

We can now combine questions (A) and (B) into the following problem P:

$$\text{(1)} \qquad \underset{z}{\text{Minimize}}\ h(\underset{\sim}{z}) \equiv \sum_{n=1}^{N} \lambda_n g_n(\underset{\sim}{z})$$

subject to

$$\text{(2)} \qquad f_m(\underset{\sim}{z}) \geqslant c_m \qquad m = 1, \ldots, M;$$

$$\text{(3)} \qquad g_n(\underset{\sim}{z}) \leqslant d_n \qquad n = 1, \ldots, N;$$

$$\text{(4)} \qquad \sum_{j=1}^{J} z_{i,j} \leqslant w_i \qquad i = 1, \ldots, I;$$

$$\text{(5)} \qquad z_{i,j} \in Z^+ \qquad i = 1, \ldots, I;\ j = 1, \ldots, J;$$

where Z^+ is the set of nonnegative integers. If problem P is infeasible, then the answer to question (A) is clearly "no"; otherwise, an answer to question (B) is ensured because the number of allocations which satisfies constraints (4) and (5) is finite.

3. AN ALGORITHM

Problem P admits solution by implicit enumeration. The following algorithm is based upon the lexicographic technique of Lawler and Bell, [4] though, unlike the Lawler-Bell approach, this algorithm does not use binary vectors. We first identify the matrix $\underset{\sim}{z}$ with a vector $\underset{\sim}{\hat{z}}$. This can be done in a number of ways, one of which is through the following relationship:

$$\text{(6)} \qquad \hat{z}_k = z_{i,j}, \qquad k = i + (j-1) \cdot I;\ i = 1, \ldots, I;\ j = 1, \ldots, J.$$

Note that this can be reversed as follows:

$$z_{i,j} = \hat{z}_k,\ i = k - \left\langle \frac{k-1}{I} \right\rangle \cdot I;\ j = \left\langle \frac{k-1}{I} \right\rangle + 1;\ k = 1, \ldots, K = I \cdot J,$$

where $\langle x \rangle$ is the largest integer less than or equal to x. With this in mind, we will drop the circumflex, and in the discussion that follows, all allocations will be vectors in Z_K^+, i.e., K-dimensional vectors with nonnegative integer components. We require two binary relations between vectors in Z_K^+:

COMPONENTWISE (PARTIAL) ORDERING: We write $\underset{\sim}{x} \geqslant \underset{\sim}{y}$ if $x_k \geqslant y_k$, $k = 1, \ldots, K$; $\underset{\sim}{x} > \underset{\sim}{y}$ if $\underset{\sim}{x} \geqslant \underset{\sim}{y}$ and $x_k > y_k$ for at least one k.

LEXICOGRAPHIC ORDERING: We write $\underset{\sim}{x} \underset{L}{>} \underset{\sim}{y}$ if $x_{k'} > y_{k'}$ where $k' = \max_{1 \leqslant k \leqslant K} \{k \mid x_k \neq y_k\}$, and $\underset{\sim}{x} \underset{L}{\geqslant} \underset{\sim}{y}$ if $\underset{\sim}{x} \underset{L}{>} \underset{\sim}{y}$ or $\underset{\sim}{x} = \underset{\sim}{y}$.

Let

$$\mathcal{G} = \left\{ \underset{\sim}{z} \in Z_K^+ \mid z_k \leqslant w_h \text{ for } k = 1, \ldots, K;\ h = k - \left\langle \frac{k-1}{I} \right\rangle \cdot I \right\}.$$

Thus $\mathcal{G}$ is a set of allocations that satisfies constraint (5) of problem P, and clearly contains all allocations that satisfy constraint (4), and so must contain all solutions to problem P, providing problem P is feasible. Since $\underset{L}{\geqslant}$ totally orders $\mathcal{G}$, we could enumerate all the points of $\mathcal{G}$ and find the solution to P in this manner. However, the monotonicity of the objective and constraint functions will permit us to skip over many infeasible and/or nonoptimal points. To see this, we need some notation. Consider a vector $\underset{\sim}{z} \in \mathcal{G}$. We will denote by $\underset{\sim}{z} + 1$, the vector $\underset{\sim}{x}$, if it exists, satisfying

$$(7) \qquad \begin{cases} \underset{\sim}{x} \in \mathcal{G} \\ \underset{\sim}{x} \underset{L}{>} \underset{\sim}{z} \\ \underset{\sim}{y} \underset{L}{>} \underset{\sim}{z} \rightarrow \underset{\sim}{y} \underset{L}{\geqslant} \underset{\sim}{x}. \end{cases}$$

At most one such vector exists, but may fail to exist because of the boundedness of G. The vector $\underset{\sim}{z} - 1$ will be that vector $\underset{\sim}{x}$, if it exists, satisfying

$$(8) \qquad \begin{cases} \underset{\sim}{x} \in \mathcal{G} \\ \underset{\sim}{z} \underset{L}{>} \underset{\sim}{x} \\ \underset{\sim}{z} \underset{L}{>} \underset{\sim}{y} \rightarrow \underset{\sim}{x} \underset{L}{>} \underset{\sim}{y} \end{cases}$$

This vector will always exist except for $\underset{\sim}{z} \equiv 0$. The vector $\underset{\sim}{z}^*$ will be that $\underset{\sim}{x}$, if it exists, satisfying

$$(9) \qquad \begin{cases} \underset{\sim}{x} \in \mathcal{G} \\ \underset{\sim}{x} \underset{L}{>} \underset{\sim}{z} \\ \underset{\sim}{x} \not\geqslant \underset{\sim}{z} \\ \left[(\underset{\sim}{y} \underset{L}{>} \underset{\sim}{z}) \wedge (\underset{\sim}{y} \not\geqslant \underset{\sim}{z}) \right] \rightarrow \underset{\sim}{y} \underset{L}{\geqslant} \underset{\sim}{x} \end{cases}$$

Intuitively, $\underset{\sim}{z}^*$ is the first vector in $\mathcal{G}$ following $\underset{\sim}{z}$ (in the lexicographic ordering) which is not (componentwise) greater than or equal to $\underset{\sim}{z}$. For some $\underset{\sim}{z}$, $\underset{\sim}{z}^*$ may not exist; however, we will adopt the following convention: For any $\underset{\sim}{z}$ for which $\underset{\sim}{z}^*$ does not exist, we will set

$$(\underset{\sim}{z}^* - 1)_k = w_h \text{ for } h = k - \left\langle \frac{k-1}{I} \right\rangle \cdot I,\ k = 1, \ldots, K,$$

thereby ensuring that $\underset{\sim}{z}^* - 1$ exists for every $\underset{\sim}{z} \in \mathcal{G}$. Crucial to the algorithm is the observation that for any $\underset{\sim}{z} \in \mathcal{G}$, any $\underset{\sim}{y}$ that satisfies $\underset{\sim}{z} \underset{L}{\leqslant} \underset{\sim}{y} \underset{L}{\leqslant} \underset{\sim}{z}^* - 1$ also satisfies $\underset{\sim}{y} \geqslant \underset{\sim}{z}$.

Figure 1 outlines the fundamentals of the algorithm. A brief inspection of the flow chart will make clear that the algorithm must terminate after a finite number of steps. If $\bar{h} = \infty$ upon termination, the problem is infeasible, otherwise an optimum integer allocation will always be found. The order in which the constraints are examined was chosen because, for certain applications, this order was efficient. However, we make no claim that this is, in any sense, an optimal ordering. For other applications, a different sequence of constraint evaluations might well proved to be better.

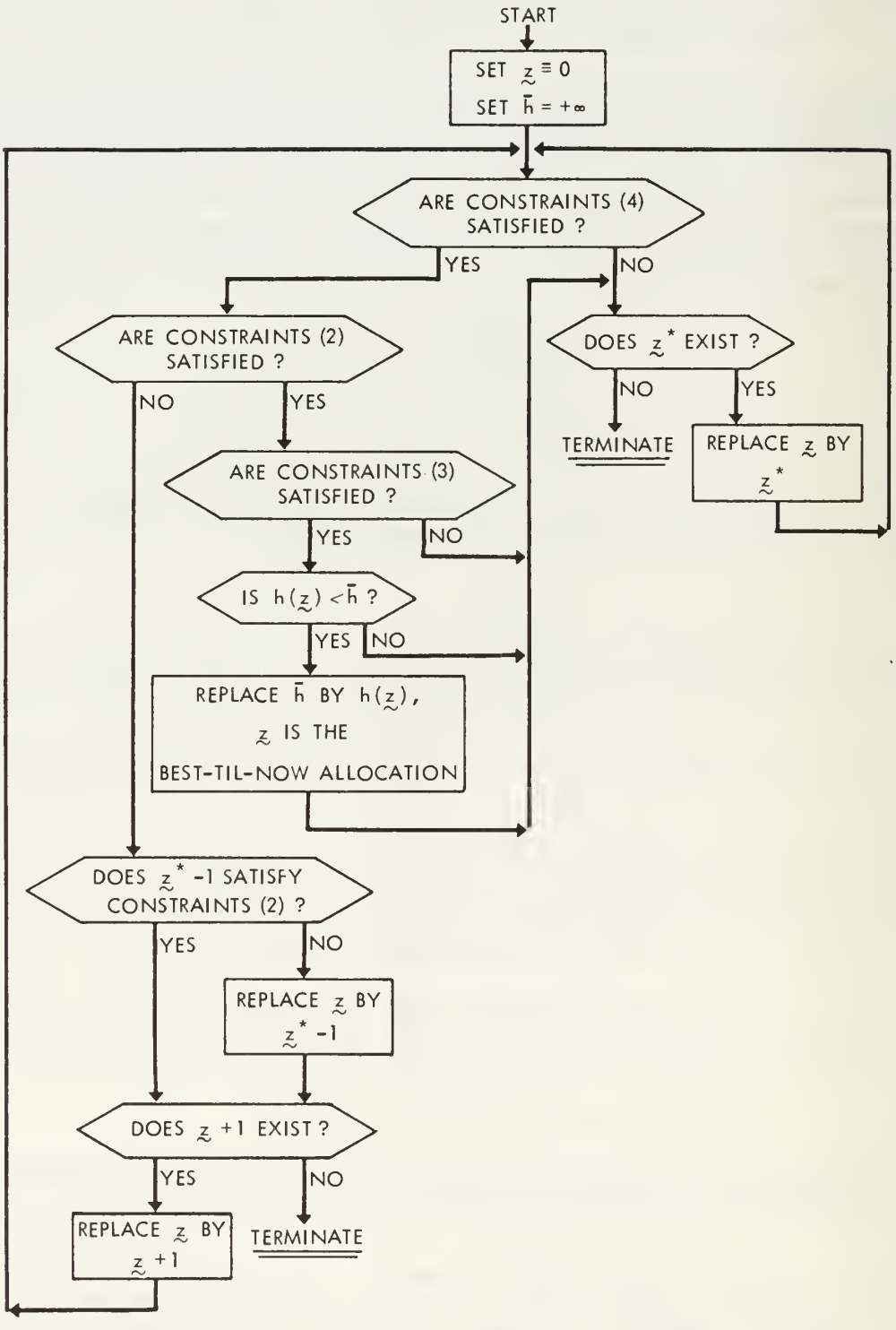


FIGURE 1. An implicit enumeration algorithm.

4. A CLASS OF EXAMPLES

We will now look at a class of examples with point targets and nontargets, where the destruction of any target or nontarget is a binomial random variable with probability of kill dependent on the allocation, but with independent weapons effects. We will use Cartesian coordinates to specify locations; in particular, target coordinates are (x_m, y_m), $m = 1, \ldots, M$; nontarget coordinates are (μ_n, ν_n), $n = 1, \ldots, N$; and aimpoint coordinates are (ξ_j, ζ_j), $j = 1, \ldots, J$. For response functions we will employ "probability of kill" which is computed as follows: Let $p_{i,j}^m$ be the probability that a single weapon of type i, allocated to aimpoint j, destroys target m, conditioned on the weapon's arrival at its destination. The probability that a type-i weapon arrives at its destination, i.e., its "reliability," is given by ρ_i. Because we have assumed independence of weapon effects, it is not difficult to compute the total probability that target m is destroyed by allocation z, which is

$$f_m(\underset{\sim}{z}) = 1 - \prod_{i=1}^{I} \prod_{j=1}^{J} (1 - \rho_i p_{i,j}^m)^{z_{i,j}}.$$

Similarly, we denote by $p_{i,j}^n$ the conditional probability that a single type-i weapon allocated to aimpoint j destroys nontarget n. Therefore, the probability that allocation $\underset{\sim}{z}$ destroys nontarget n is

$$g_n(\underset{\sim}{z}) = 1 - \prod_{i=1}^{I} \prod_{j=1}^{J} (1 - \rho_i p_{i,j}^n)^{z_{i,j}}.$$

Although the values of the parameters $\{p_{i,j}^m\}$ and $\{p_{i,j}^n\}$ can be entirely arbitrary, within the obvious limits

$$0 \leqslant p_{i,j}^m \leqslant 1 \qquad m = 1, \ldots, M;\ i = 1, \ldots, I;\ j = 1, \ldots, J,$$

$$0 \leqslant p_{i,j}^n \leqslant 1 \qquad n = 1, \ldots, N;\ i = 1, \ldots, I;\ j = 1, \ldots, J,$$

we will use, for tutorial purposes, the following formulae, which are not unreasonable approximations to certain types of weapon damage curves and have been proposed by other analysts (see, for example, Eckler [1], or McNolty [5]):

$$p_{i,j}^m = \exp\{-\alpha_{i,m}[(x_m - \xi_j)^2 + (y_m - \zeta_j)^2]\} \quad m = 1, \ldots, M;\ i = 1, \ldots, I;\ j = 1, \ldots, J,$$

$$p_{i,j}^n = \exp\{-\beta_{i,n}[(\mu_n - \xi_j)^2 + (\nu_n - \zeta_j)^2]\} \quad n = 1, \ldots, N;\ i = 1, \ldots, I;\ j = 1, \ldots, J, \tag{10}$$

where all $\alpha_{i,m}$, $\beta_{i,n}$ are nonnegative real numbers. The parameters $\{\alpha_{i,m}\}$ and $\{\beta_{i,n}\}$ are measures of the rate at which weapon effects decrease with distance.

With these conventions, we can now write explicitly the problem P' which comprises this class of examples:

P': Given nonnegative weights λ_n, $n = 1, \ldots, N$, and the values of

$c_m \epsilon [0, 1]$,	$m = 1, \ldots, M$
$d_n \epsilon [0, 1]$,	$n = 1, \ldots, N$
$w_i \epsilon Z^+$,	$i = 1, \ldots, I$
$\rho_i \epsilon [0, 1]$,	$i = 1, \ldots, I$
$\alpha_{i,m} \geqslant 0$,	$i = 1, \ldots, I;\ m = 1, \ldots, M$
$\beta_{i,n} \geqslant 0$,	$i = 1, \ldots, I;\ n = 1, \ldots, N$
x_m, y_m,	$m = 1, \ldots, M$
μ_n, ν_n,	$n = 1, \ldots, N$
ξ_j, ζ_j,	$j = 1, \ldots, J$

$$\underset{z}{\text{minimize}}\ h(\underline{z}) = \sum_{n=1}^{N} \lambda_n \left\{1 - \prod_{i=1}^{I}\prod_{j=1}^{J}\left[1 - \rho_i \exp\{-\beta_{i,n}[(\mu_n - \xi_j)^2 + (\nu_n - \zeta_j)^2]\}\right]^{z_{i,j}}\right\}$$

subject to

$$f_m(\underline{z}) =$$

$$1 - \prod_{i=1}^{I}\prod_{j=1}^{J}(1 - \rho_i \exp\{-\alpha_{i,m}[(x_m - \xi_j)^2 + (y_m - \zeta_j)^2]\})^{z_{i,j}} \geqslant c_m \quad m = 1,\ldots, M$$

$$g_n(\underline{z}) =$$

$$1 - \prod_{i=1}^{I}\prod_{j=1}^{J}(1 - \rho_i \exp\{-\beta_{i,n}[(\mu_n - \xi_j)^2 + (\nu_n - \zeta_j)^2]\})^{z_{i,j}} \leqslant d_n \quad n = 1, \ldots, N$$

$$\sum_{j=1}^{J} z_{i,j} \leqslant w_i \qquad i = 1, \ldots, I$$

$$z_{i,j} \in Z^+ \qquad i = 1, \ldots, I;\ j = 1, \ldots, J.$$

5. COMPUTER APPLICATIONS

A FORTRAN routine to solve problems of the type given by P' was written for the CDC 6400 computer, and was used to solve the numerical example of this section. The values of the parameters are listed in Tables 1-6. The configuration of the targets, nontargets, and aimpoints is depicted in Fig. 2.

TABLE 1. *Target Parameters*

$M = 2$

m	x_m	y_m	c_m
1	−1	0	0.8
2	1	0	0.8

TABLE 2. *Nontarget Parameters*

$N = 4$

n	μ_n	ν_n	λ_n	d_n
1	−2	0	2	0.3
2	−1	−1	4	0.3
3	1	1	6	0.3
4	2	0	8	0.3

TABLE 3. *Aimpoint Parameters*

$J = 5$

j	ξ_j	ζ_j
1	−1	1
2	−1	0
3	0	0
4	1	0
5	1	−1

TABLE 4. *Weapon Parameters*

$I = 2$

i	w_i	ρ_i
1	6	0.9
2	6	0.7

TABLE 5. *Components of* $\underset{\sim}{\alpha}$

i \ m	1	2
1	0.1	0.1
2	0.5	0.5

TABLE 6. *Components of* $\underset{\sim}{\beta}$

i \ n	1	2	3	4
1	0.05	0.1	0.1	0.09
2	0.80	0.8	0.8	0.80

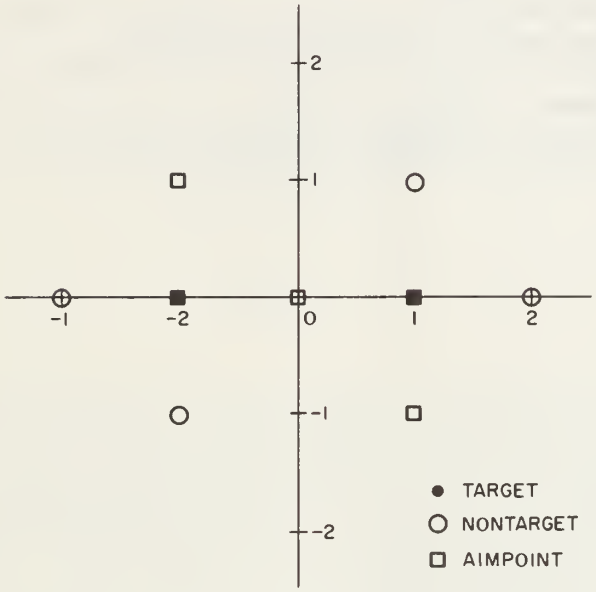


FIGURE 2. Configuration of the example.

The routine ran for five seconds to compute the optimal solution $\underset{\sim}{\hat{z}}$ given in Table 7. It is interesting to note that if all the d_i are changed to 1.0, which is equivalent to removing the individual nontarget damage constraints, then the optimal allocation is $\underset{\sim}{z'}$, given in Table 8. In this latter case, we have reduced total collateral damage over that given in Table 7, but only at the expense of considerably greater damage to two of the nontargets.

TABLE 7. *Optimal Allocation* $\underset{\sim}{\hat{z}}$

i \ j	1	2	3	4	5
1	0	0	0	0	0
2	2	0	1	0	2

$h(\underset{\sim}{\hat{z}}) = 5.2$ $\quad g_2(\underset{\sim}{\hat{z}}) = 0.24$
$f_1(\underset{\sim}{\hat{z}}) = 0.83$ $\quad g_3(\underset{\sim}{\hat{z}}) = 0.24$
$f_2(\underset{\sim}{\hat{z}}) = 0.83$ $\quad g_4(\underset{\sim}{\hat{z}}) = 0.28$
$g_1(\underset{\sim}{\hat{z}}) = 0.28$

TABLE 8. *Optimal Allocation* $\underset{\sim}{z'}$

	1	2	3	4	5
1	0	0	0	0	0
2	0	0	3	0	0

$h(\underset{\sim}{z'}) = 4.5$ $\quad g_2(\underset{\sim}{z'}) = 0.37$
$f_1(\underset{\sim}{z'}) = 0.81$ $\quad g_3(\underset{\sim}{z'}) = 0.37$
$f_2(\underset{\sim}{z'}) = 0.82$ $\quad g_4(\underset{\sim}{z'}) = 0.08$
$g_1(\underset{\sim}{z'}) = 0.08$

BIBLIOGRAPHY

[1] Eckler, A. R. and S. A. Burr, "Mathematical Models of Target Coverage and Missile Allocation," Military Operations Research Society, Alexandria, Va. (1972).

[2] Grotte, J. H., "A Nontarget Damage Minimizer (NDM) Which Achieves Targeting Objectives," Paper P-1106 (also WSEG Report 256, Weapons Systems Evaluation Group, Arlington, Va.). Institute for Defense Analyses, Arlington, Va. (Nov. 1975). (Vol. I, Unclassified; Vol. II, Secret).

[3] Grotte, J.H., "A Targeting Model that Minimizes Collateral Damage," Paper P-1135, Institute for Defense Analyses, Arlington, Va., (Dec. 1975).

[4] Lawler, E. L. and M. D. Bell, "A Method for Solving Discrete Optimization Problems," Operations Research *14,* 1098-1112 (1966).

[5] McNolty, F., "Expected Coverage for Targets of Nonuniform Density," Operations Research *16*, 1027-1040 (1968).

DIFFERENTIAL-GAME EXAMINATION OF OPTIMAL TIME-SEQUENTIAL FIRE-SUPPORT STRATEGIES

James G. Taylor*

Department of Operations Research
Naval Postgraduate School
Monterey, California

ABSTRACT

Optimal time-sequential fire-support strategies are studied through a two-person zero-sum deterministic differential game with closed-loop (or feedback) strategies. Lanchester-type equations of warfare are used in this work. In addition to the max-min principle, the theory of singular extremals is required to solve this prescribed-duration combat problem. The combat is between two heterogeneous forces, each composed of infantry and a supporting weapon system (artillery). In contrast to previous work reported in the literature, the attrition structure of the problem at hand leads to force-level-dependent optimal fire-support strategies with the attacker's optimal fire-support strategy requiring him to sometimes split his artillery fire between enemy infantry and artillery (counterbattery fire). A solution phenomnon not previously encountered in Lanchester-type differential games is that the adjoint variables may be discontinuous across a manifold of discontinuity for both players' strategies. This makes the synthesis of optimal strategies particularly difficult. Numerical examples are given.

1. INTRODUCTION

The allocation of a specific weapon system type to an acquired target is an important tactical decision in the fire-support process.** Accordingly, the determination of optimal (or even good) fire-distribution strategies for supporting weapon systems† is a major problem of military operations research. The problem is of interest to the military tactician because he may need a clearer understanding of the circumstances under which a supporting weapon system (such as artillery) should engage the enemy's primary weapon system (i.e. infantry) and when it should engage the enemy's supporting weapon systems.

*This research was supported by the Office of Naval Research (partially through the Foundation Research Program at the Naval Postgraduate School and partially through direct funding).

**See pp. I-33 to I-43 of Ref. [26] for a further discussion.

†See Ref. [38] for a brief discussion of the distinction between a "primary" weapon system (or infantry) and a "supporting" weapon system.

In this paper we will examine the dependence of optimal time-sequential fire-support strategies on the form of the combat attrition model. Previous work by Weiss [38] and Kawara [22] suggests that an optimal fire-support strategy consists in always concentrating all fire on one enemy target type (although this target type may change over time). We will consider a differential game with slightly different combat dynamics than the fire-support differential game recently considered by Kawara [22] and show that optimal fire-support strategies quite different in structure than those obtained by him may arise. Moreover, the solution to the problem which we consider in this paper involves a solution phenomenon not previously encountered in a Lanchester-type differential game*: the dual (or adjoint) variables may be discontinuous across a manifold of discontinuity for both players' strategies.

Fire-support operations (as are any combat operations) are a complex random process [26]. We will nevertheless consider a simplified deterministic Lanchester-type model in order to develop insights into the structure of optimal time-sequential fire-support strategies. H.K. Weiss [38] has emphasized that such a model of an idealized combat situation is particularly valuable when it leads to a clearer understanding of significant relationships which would tend to be obscured in a more complex model.

The problem of determining an appropriate mixture of tactical and strategic forces (another aspect of the optimal fire-support strategy problem) was extensively debated by experts during World War II. Some analysis details may be found in the classic book by Morse and Kimball (see pp. 73-77 of [27]). The problem was studied at RAND in the late 1940's and early 1950's [16] and elsewhere [1]. It would probably not be too far-fetched to claim that this problem stimulated early research on both dynamic programming [2] and differential games [16,20]. Today the problem of determining optimal air-war strategies is being extensively studied by a number of organizations (see, for example, Refs. [17,25,29,36]). An idealized version of A. Mengel's problem [16] appears in Isaacs' book as the "War of Attrition and Attack" (see pp. 96-105 of Ref. [21]). Discrete-time versions of this problem of determining optimal "air-war" strategies have been considered by a number of workers as time-sequential combat games [5,6,15] (see also Refs. [7,13]). A related problem has been considered by Weiss [38] (see also Ref. [37]), who studied the optimal selection of targets for engagement by a supporting weapon system.** More recently, Kawara [22] has studied optimal time-sequential strategies for supporting weapon systems in an attack scenario version of Weiss' problem. Other recent work has considered various conceptual and computational aspects of time-sequential combat games [28-30].

Since our work here may be considered to be an elaboration upon and extension of Kawaras's fire-support differential game [22], we will review his main results and relate our work here to them. Kawara [22] considers combat between two heterogeneous forces, each composed of infantry (the primary weapon system) and artillery (the supporting weapon system). The time-sequential decision problem is to determine each side's optimal strategy for distributing its supporting weapon system's fire over enemy target types according to the criterion of the infantry force ratio at the prescribed-duration attack's end. Kawara concludes that each side's optimal† strategy is to always concentrate all supporting fire on the enemy's supporting weapon system (counterbattery fire) during the early stages of battle (if the total

*We refer to a differential game as being a Lanchester-type differential game when the system dynamics are described by Lanchester-type equations of warfare [34].

**See Ref. [33], however, for a justification of the optimality of strategies given by Weiss [38]. A general solution algorithm is also presented in this paper [33].

†Kawara does not determine the optimality of extremal strategies determined for his problem (i.e. show that sufficient conditions of optimality are satisfied [4]). We use the word "extremal" to denote a trajectory on which the necessary conditions are satisfied.

prescribed length of battle is long enough) and then later to switch to concentration of all fire on the enemy's infantry. He states that this switching time "does not depend on the current strength of either side but only one the effectivenesses of both sides' supporting units" (see p. 951 of Ref. [22]). Moreover, an optimal strategy has the property of always requiring concentration of all supporting fire on enemy infantry during the final stages of battle.

Thus, Kawara concludes that for his model the optimal fire-support strategies do not depend on force levels. However, this is only true provided that the appropriate side's (in Kawara's numerical example, the defender) supporting weapon system is not reduced to a zero force level before a critical time.* Let us assume, therefore, that neither side's supporting weapon systems can be reduced to a zero force level.** For this condition the optimal fire-support strategies *are* force-level independent and may be expressed solely in terms of "time-to-go" in the prescribed duration battle. The purpose of this paper is to show that a tactically realistic variation in the attrition equations leads to a problem with force-level-dependent optimal fire-support strategies. This result has an important implication for tactical decision-making: optimal time-sequential allocation of fire-support resources depends not only on initial intelligence estimates but also on a continuous monitoring of the evolution of the course of combat.

Thus, the purpose of this paper is to illustrate the dependence of optimal fire-support strategies on the nature of Lanchester-type combat attrition equations [34]. We consider a slight variation in Kawara's problem (i.e. different combat dynamics) for which the structure of the optimal strategy of one of the combatants is significantly different than that in the original problem [22]: the optimal strategy of one combatant depends directly upon the enemy's force levels and is no longer to always concentrate all fire on either the enemy's primary or supporting weapon system. Furthermore, we will show that an optimal strategy in which a side divides the fire of its supporting weapon system between the enemy's primary (infantry) and supporting systems *can only* occur when the enemy's infantry has some fire effectiveness (in the sense of a nonzero Lanchester attrition-rate coefficient) against his infantry. The optimal strategy of one side to sometimes split its fire is very similar to that which occurs in a one-sided (optimal control) problem previously considered by us [31] (see also Ref. [32]) for the optimal distribution of fire by a homogeneous force in combat against heterogeneous forces. In Ref. [31], the enemy consisted of two weapon-system types, each of which undergoes attrition at a rate proportional to the product of the numbers of firers and targets (referred to, for convenience, as "linear-law" attrition). In fact, this previous work of ours [31] was the motivation for our examination here of other attrition structures in Kawara's problem.

2. KAWARA'S FIRE-SUPPORT DIFFERENTIAL GAME

Since Kawara's fire-support differential game is the point of departure for this paper, we will review the development of his model. The reader will find it convenient to compare the mathematical statement of Kawara's problem (1) of this paper, with the fire-support differential game studied in this paper, equations (2), in order to understand the dependence of optimal fire-support strategies on the mathematical form of the attrition equations.

Kawara [22] considers the attack of heterogeneous X forces against the static defense of heterogeneous Y forces. Both the X and Y forces are composed of two types of units: primary units (or infantry) and fire-support units (or artillery). The X (infantry denoted as X_1) launches an attack against the Y infantry (denoted as Y_1). We consider that phase of the attack

*See the expression for T_2 on p. 949 of Ref. [22] and its plot in Fig. 4 of Ref. [22].

**Initial force levels and the known length of battle may be sufficient to guarantee this for a given set of (or range of values of) Lanchester attrition-rate coefficients.

which may be called the "approach to contact." This is the time from the initiation of the advance of the X_1 forces towards the Y_1 defensive position until the X_1 forces actually make contact with the enemy infantry. It is assumed that this time is fixed and known to both sides and that infantry fire has negligible effectiveness against the enemy's infantry during this time. During this time the fire-support units remain stationary, and each unit has the capability to deliver either "point-fire" counterbattery fire against enemy artillery or "area fire" against the enemy's infantry.

It is the objective of each side to attain the most favorable infantry force ratio* possible at the end of the "approach to contact," at which time the force separation between the opposing infantries is zero and artillery fires must be lifted from the enemy's infantry in order not to kill friendly forces. Thus, the decision problem facing each commander is to determine the "best" distribution of artillery fire over time between enemy infantry and enemy artillery in order to maximize the quotient of friendly infantry (numerical) strength divided by enemy infantry strength at the end of the approach to contact. This situation is shown diagrammatically in Fig. 1. The reader is referred to Kawara's paper [22] for further details of the model's development. It should be pointed out that this model also applies to the case of an amphibious landing and the determination of the optimal time-sequential allocation of the supporting fires of naval ship guns.

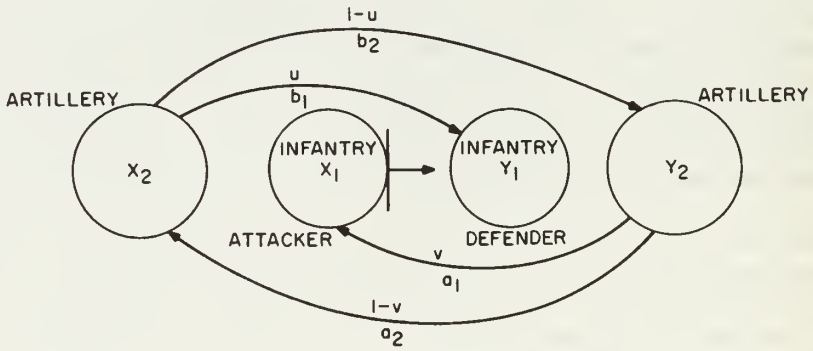


FIGURE 1. Diagram of Kawara's Fire-Support Differential Game.

Mathematically, the problem may be stated as the following.**

$$\underset{U}{\text{maximize}}\ \underset{V}{\text{minimize}} \left\{ \frac{x_1(t_f)}{y_1(t_f)} \right\},$$

$$\text{with stopping rule: } t_f - T = 0,$$

$$\text{subject to: (battle dynamics)} \quad \frac{dx_1}{dt} = -va_1x_1y_2,$$

$$\frac{dy_1}{dt} = -(1-v)a_2y_2, \tag{1}$$

$$\frac{dy_1}{dt} = -ub_1y_1x_2,$$

$$\frac{dy_2}{dt} = -(1-u)b_2x_2,$$

*See Ref [35] for some insights into the dynamics of combat from considering the force ratio.

**We use capital letters to denote the closed-loop (or feedback) strategies [19] of the players and the corresponding lower case letters to denote the corresponding strategic variables [4]. A strategic variable is the realization (or outcome) of a strategy. Thus, $u(t) = U(t,\mathbf{x},\mathbf{y})$ and $v(t) = V(t,\mathbf{x},\mathbf{y})$.

with initial conditions

$$x_i(t = 0) = x_i^o \text{ and } y_i(t = 0) = y_i^o \text{ for } i = 1, 2,$$

and

$$x_1, x_2, y_1, y_2 \geqslant 0 \text{ (State Variable Inequality Constraints)},$$
$$0 \leqslant u, v \leqslant 1 \text{ (Strategic Variable Inequality Constraints)},$$

where

$x_1(t)$ is the number of X infantry (i.e. X_1) at time t,
$x_2(t)$ is the number of X artillery (i.e. X_2) at time t,
similarly for $y_1(t)$ and $y_2(t)$,

a_i is a constant (Lanchester) attriction-rate coefficient* (reflecting the effectiveness of Y_2 fire against X_i), similarly for b_i,

and

$u(v)$ is the fraction of $X(Y)$ artillery fire directed at opposing infantry forces.

We observe that for $T < +\infty$ it follows from the battle dynamics (1) that $x_1(t), y_1(t) > 0$ for all $t \in [0, T]$. Thus, the only State Variable Inequality Constraints (SVIC's) that must be considered are $x_2, y_2 \geqslant 0$.

Kawara's results and conclusions [22] have been previously discussed in Section 1.

3. ANOTHER MODEL FOR OPTIMAL FIRE-SUPPORT ALLOCATION

In this paper we will study a variation of Kawara's [22] fire-support differential game (1) just given. We will see that for this problem the structure of the optimal fire-support strategy for the attacker has a fundamentally different form than that for (1): the attacker must sometimes split his fire between the defender's primary supporting units in order to "avoid overkill." Furthermore, the nature of this split in an optimal strategy depends on the allocation of the defender's supporting fires.

Let us again consider the attack of heterogeneous X forces against the static defense of heterogeneous Y forces. Each side is composed of primary units (or infantry) and fire-support units (or artillery). The X infantry (denoted as X_1) launches an attack against the position held by the Y infantry (denoted as Y_1). Again, we will consider only the "approach to contact" phase of the battle. This phase is the time from the initiation of the advance of the X_1 forces towards the Y_1 defensive position until the X_1 forces actually make contact with the enemy infantry. It is assumed that this time is fixed and known to both sides.

Using "cover and concealment," the X_1 forces begin their advance against the Y_1 forces from a distance and move towards the Y_1 position. The objective of the X_1 forces during the "approach to contact" is to close with the enemy position as rapidly as possible. Accordingly, small-arms fire by the X_1 forces is held at a minimum, or firing is done "on the move" to facilitate their rapid movement. It is not unreasonable, therefore, to assume that the effectiveness of X_1's fire "on the move" is negligible against Y_1. We assume, however, that the defensive Y_1 fire causes attrition to the advancing X_1 forces at a rate proportional to the product of the numbers of firers and targets. Two possible justifications of this are as follows: because of the

*See Ref. [10] (also [8,9]) for methodology for the prediction of such coefficients from weapon system performance data.

movement (and intermittent concealment) of the X_1 forces and the distance involved, the Y_1 defenders either (a) fire into a constant (but moving) area without precise knowledge of the consequences of their fire, or (b) fire at X_1 targets, with the result that the time to acquire such a target is inversely proportional to the density of X_1 forces and much greater than the time to kill an acquired target. Under each of these sets of circumstances the assumed form of attrition has been hypothesized to occur [11,37].

During the "approach to contact," the fire-support units remain stationary. Each unit has the capability to deliver counterbattery fire against enemy artillery or "area fire" against the enemy's infantry. In other words, we assume that each side's fire-support units fire into the (constant) area containing the enemy's infantry without feedback as to the destructiveness of this fire. On the other hand, the effectiveness of counterbattery fire is not symmetric with respect to the two combatants. We assume that the defender has the capability (for example, through the use of aerial observers) to sense when an enemy supporting unit has been destroyed so that fire may be immediately shifted to a new target*, and that fire is uniformly distributed over the survivors.† The attacker, however, either (a) does not have the capability to sense destruction of enemy fire-support units accurately (and hence distributes his fire uniformly over the (constant) area occupied by the defender's fire-support units), or (b) does have adequate fire assessment capability, with the result that target acquisition times (which are inversely proportional to the density of the enemy's fire-support units) are much larger than the time to destroy an acquired target. This leads to a Y_2 attrition rate proportional to the product of the numbers of X_2 firers and Y_2 targets [11,37].

It is the objective of each side to attain the most favorable infantry force ratio possible at the end of the "approach to contact," at which time the force separation between the opposing infantries is zero and artillery fires must be lifted from the enemy infantry's position in order not to also kill friendly forces. Thus, the decision problem facing each side is to determine the "best" distribution of artillery fire between enemy infantry and artillery over time in order to maximize the infantry force ratio at the time of contact between the two infantry forces. This situation is shown diagrammatically in Fig. 2.

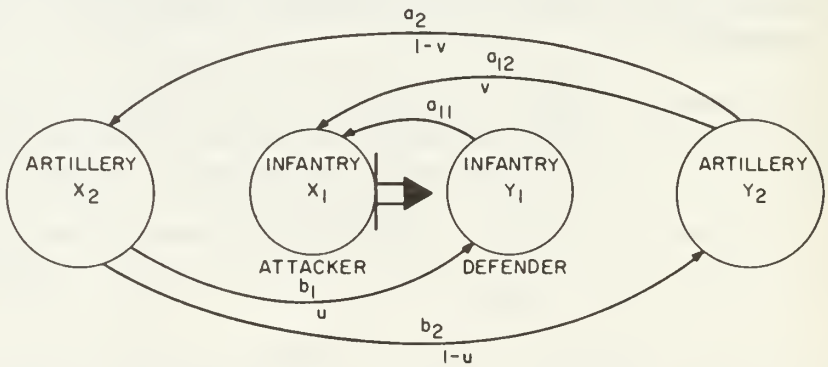


FIGURE 2. Diagram of Fire-Support Differential Game studies in this paper.

*Alternatively, we may think that the attacker has massed so much artillery that X_2 targets are always easily acquired by Y_2 once an Y_2 unit has been destroyed. Moreover, it will be assumed below that the initial X_2 force level is sufficiently large to guarantee that it is never driven to zero.

†This assumption is not essential for the structure of X_2's optimal fire-support stragety. A similar structural result may be obtained when X_2's attrition is the same form as that for Y_2. We have made the above assumption, moreover, so that the resultant attrition model is most similar to Kawara's [22] but yet yields significantly different results for the attacker's fire-support strategy.

The above assumptions lead to the following differential game with an attrition structure slightly different from that in Kawara's problem [22].

$$\underset{U}{\text{maximize}}\ \underset{V}{\text{minimize}} \left\{ \frac{x_1(t_f)}{y_1(t_f)} \right\},$$

with stopping rule: $t_f - T = 0$,

$$\text{(2)} \qquad \underset{\text{(battle dynamics)}}{\text{subject to:}} \quad \frac{dx_1}{dt} = -a_{11}x_1y_1 - va_{12}x_1y_2,$$

$$\frac{dx_2}{dt} = -(1-v)a_2y_2,$$

$$\frac{dy_1}{dt} = -ub_1y_1x_2,$$

$$\frac{dy_2}{dt} = -(1-u)b_2y_2x_2,$$

with initial conditions

$$x_i(t=0) = x_i^o \text{ and } y_i(t=0) = y_i^o \text{ for } i = 1,2,$$

and

$$x_1,x_2,y_1,y_2 \geqslant \text{(State Variable Inequality Constraints)},$$

$$0 \leqslant u,v \leqslant 1 \text{ (Strategic Variable Inequality Constraints)},$$

where all symbols are (essentially) the same as defined above for problem (1).

We observe that for $T < +\infty$ it follows from the battle dynamics (1) that $x_1(t)$, $y_1(t)$, and $y_2(t) > 0$ for all $t \in [0,T]$. Thus, the only SVIC that must be considered is $x_2 \geqslant 0$. However, let us *assume that the force level of the attacker's artillery is never reduced to zero.* In other words, we consider the special case in which x_2^o and T are such that $x_2(T) > 0$.

4. CHARACTERIZATION OF OPTIMAL FIRE-DISTRIBUTION STRATEGIES FOR THE SUPPORTING WEAPON SYSTEMS

It should be clear that in (2) above we have $a_{11},a_{12}, a_2,b_1,b_2 > 0$. Although the results of A. Friedman [14] concerning existence of value do not apply directly to our fire-support differential game (2), they do apply to a suitably modified version. If we were to consider a version of this problem with $\frac{dx_2}{dt} = -(1-v)a_2y_2 + r_2$ where $r_2 > 0$, then it may be shown (see pp. 210-230 of Ref. [14]) that this "modified" differential game has value and that a saddle point exists in pure strategies (see pp. 234-235 of Ref. [14]). We will now develop the basic necessary conditions of optimality for (2).

For $x_1,x_2,y_1,y_2 > 0$, the Hamiltonian for (2) is given by [12]

$$\text{(3)} \qquad \begin{aligned} H(t,\mathbf{x},\mathbf{y},\mathbf{p},\mathbf{q},u,v) &= p_1(a_{11}x_1y_1 + va_{12}x_1y_2) - p_2a_2(1-v)y_2 \\ &\quad - q_1ub_1y_1x_2 - q_2(1-u)b_2y_2x_2, \end{aligned}$$

where we have adopted the following correspondence between state and adjoint variables:

state Variable	dual variable	
x_i	p_i	
y_i	q_i	for $i = 1,2$.

The adjoint system of differential equations for the dual variables is

$$\frac{dp_1}{dt} = -\frac{\partial H}{\partial x_1} = a_{11}y_1p_1 + v^*a_{12}y_2p_1 \quad \text{with } p_1(T) = \frac{1}{y_1^f}, \tag{4}$$

$$\frac{dp_2}{dt} = -\frac{\partial H}{\partial x_2} = u^*b_1y_1q_1 + (1-u^*)b_2y_2q_2 \quad \text{with } p_2(T) = 0, \tag{5}$$

$$\frac{dq_1}{dt} = -\frac{\partial H}{\partial y_1} = a_{11}x_1p_1 + u^*b_1x_2q_1 \quad \text{with } q_1(T) = -\frac{x_1^f}{(y_1^f)^2}, \tag{6}$$

$$\frac{dq_2}{dt} = -\frac{\partial H}{\partial y_2} = v^*a_{12}x_1p_1 + (1-v^*)a_2p_2 + (1-u^*)b_2x_2q_2 \quad \text{with } q_2(T) = 0. \tag{7}$$

The results of Berkovitz [3] say that $H, \mathbf{p}(t)$, and $\mathbf{q}(t)$ are continuous functions of time except possible at manifolds of discontinuity of both U^* and V^* (see Section 4.3 below).

When $x_1,x_2,y_1,y_2 > 0$, the extremal strategic-variable pair, denoted as (u^*,v^*), is determined by the *max-min principle*. Hence, we consider

$$\underset{0 \leqslant u \leqslant 1}{\text{maximize}} \quad \underset{0 \leqslant v \leqslant 1}{\text{minimize}} \; H(t, \mathbf{x}, \mathbf{y}, \mathbf{p}, \mathbf{q}, u, v),$$

so that

$$u^*(t) = \begin{cases} 1 & \text{for } S_u(t) > 0, \\ 0 & \text{for } S_u(t) < 0, \end{cases} \tag{8}$$

where the U-switching function $S_u(t)$ is given by

$$S_u(t) = b_1(-q_1)y_1 - b_2(-q_2)y_2, \tag{9}$$

and

$$v^*(t) = \begin{cases} 1 & \text{for } S_v(t) > 0, \\ 0 & \text{for } S_v(t) < 0, \end{cases} \tag{10}$$

where the V-switching function $S_v(t)$ is given by

$$S_v(t) = a_{12}p_1x_1 - a_2p_2. \tag{11}$$

It is readily shown that

$$p_1(t)x_1(t) = \text{constant} = p_1(T)x_1(T) = \frac{x_1^f}{y_1^f}, \tag{12}$$

$$\frac{d}{dt}(q_1y_1) = a_{11}\frac{x_1^f}{y_1^f}y_1(t) > 0, \tag{13}$$

and

$$\frac{dS_v}{dt} = -a_2(1-u^*)S_u(t) - a_2b_1q_1y_1. \tag{14}$$

We must further investigate the possibility of singular subarcs (Ref. [31] or Chapter 8 of Ref. [12]). Let us first show that *it is impossible to have a V-singular subarc.* In other words, $v^*(t)$ must be 0 or 1 almost everwhere in time. The impossibility of a V-singular subarc is established by showing that $\frac{dS_v}{dt} > 0$ for all $t \in [0,T]$. It is clear that

$$(1-u^*)S_u(t) \leqslant 0 \text{ for all } t \in [0,T]. \tag{15}$$

Considering (13) and the fact that $q_1(T)y_1(T) = -\dfrac{x_1^f}{y_1^f} < 0$, we see that $q_1(t)y_1(t) < 0$ for all $t \in [0,T]$, whence follows the assertion via (14).

It is possible, however, to have a U-singular subarc on which $\dfrac{\partial H}{\partial u} = 0$ (or, equivalently, $S_u(t) = 0$) for a finite interval of time. There are two cases to be considered: (a) $v^* = 1$ and (b) $v^* = 0$.

4.1. U-Singular Subarc on Which $V^* = 1$

When $v^* = 1$, it is readily computed that

$$\frac{dS_u}{dt} = -\left(\frac{x_1^f}{y_1^f}\right)(a_{11}b_1y_1 - a_{12}b_2y_2), \tag{16}$$

and

$$\frac{d^2S_u}{dt^2} = -\left(\frac{x_1^f}{y_1^f}\right)x_2\{(a_{11}b_1y_1)u^*b_1 - (a_{12}b_2y_2)(1-u^*)b_2\}. \tag{17}$$

If we consider (9), the requirement that $\dfrac{\partial H}{\partial u} = 0$ yields the *first condition for a U-singular subarc with* $V^* = 1$:

$$b_1q_1y_1 = b_2q_2y_2. \tag{18}$$

If we consider (16) and (18), the requirement that $\dfrac{d}{dt}\left(\dfrac{\partial H}{\partial u}\right) = 0$ on a singular subarc on which $\dfrac{\partial H}{\partial u} = 0$ for a finite interval of time yields the *second condition for a U-singular subarc with* $V^* = 1$:

$$a_{11}b_1y_1 = a_{12}b_2y_2. \tag{19}$$

On a subarc on which the first and second conditions for a singular subarc hold, we additionally require that $\dfrac{d^2}{dt^2}\left(\dfrac{\partial H}{\partial u}\right) = 0$ so that (17) yields the *singular strategic-variable value* required to keep the system on the singular subarc

$$u^*(t) = \frac{b_2}{b_1+b_2}. \tag{20}$$

Checking the generalized Legendre-Clebsch condition* [23,24] $\dfrac{\partial}{\partial u}\left\{\dfrac{d^2}{dt^2}\left(\dfrac{\partial H}{\partial u}\right)\right\} \geqslant 0$, we find that on a subarc on which (18) and (19) hold we have

$$\frac{\partial}{\partial u}\left\{\frac{d^2}{dt^2}\left(\frac{\partial H}{\partial u}\right)\right\} = \left(\frac{x_1^f}{y_1^f}\right)(x_2)^2\{a_{11}(b_1)^2y_1 + a_{12}(b_2)^2y_2\} > 0.$$

We may write the equation of the U-singular surface (see p. 683 of [31]) as

$$\frac{y_1}{y_2} = \frac{a_{12}b_2}{a_{11}b_1} \quad \text{for } v^* = 1. \tag{21}$$

*This is a necessary condition for optimality. R. Isaacs [21] gives an equivalent condition [18].

4.2. U-Singular Subarc on Which V* = 0

When $v^* = 0$, it is readily computed that

$$\frac{dS_u}{dt} = -\frac{x_1^f}{y_1^f}(a_{11}b_1y_1 - a_{12}b_2y_2) - b_2y_2S_v(t), \tag{22}$$

and

$$\frac{d^2S_u}{dt^2} = -u^*b_1x_2\frac{dS_u}{dt} + a_2b_2y_2\{-u^*S_u(t) + b_2q_2y_2 + p_2x_2[u^*b_1 - (1-u^*)b_2]\}, \tag{23}$$

so that the *first and second conditions for a U-singular subarc with* $V^* = 0$ are, respectively, (18) and

$$a_{11}b_1y_1 = a_{12}b_2y_2 + b_2y_2\left(\frac{y_1^f}{x_1^f}\right)\{-S_v(t)\}. \tag{24}$$

It should be noted [18] that the above singular surface exists in $\mathbf{x} - \mathbf{p}$ space. It is convenient to write

$$\frac{y_1}{y_2} = \frac{a_{12}b_2}{a_{11}b_1} + \frac{b_2y_1^f}{a_{11}b_1x_1^f}\{-S_v(t)\} \quad \text{for } v^* = 0. \tag{25}$$

The *singular strategic-variable value* is given by

$$u^*(t) = \left(\frac{b_2}{b_1+b_2}\right)\left(1 - \frac{q_2y_2}{p_2x_2}\right). \tag{26}$$

The requirement $u^* \leqslant 1$ yields that on a U-singular subarc with $V^* = 0$ we must have

$$b_2(-q_2)y_2 \leqslant b_1p_2x_2. \tag{27}$$

It is readily checked that the generalized Legendre-Clebsch condition is satisfied.

4.3 Discontinuity of Adjoint Variables Across Manifold of Discontinuity of Both U* and V*

It is convenient to introduce the *backwards time* τ defined by

$$\tau = T - \tau. \tag{28}$$

From (20) and (26), we see that $u^*(\tau)$ must change, in general, discontinuously from $b_2/(b_1+b_2)$ to $b_2/(b_1+b_2)(1-q_2y_2/(p_2x_2))$ whenever $v^*(\tau)$ changes from 1 to 0. Let us consider the totality of trajectories on which this happens. The locus of points in the $t, \mathbf{x}, \mathbf{y}$ — space for such simultaneous switches is then a manifold of discontinuity of both U^* and V^*. Across such a manifold the adjoint variables need not be continuous [3].

Let $\tau_v = \tau_v(\mathbf{x}, \mathbf{y})$ denote the backwards time at which $v^*(\tau)$ changes from 1 to 0. For future purposes, it will be convenient to consider a simultaneous switch with u^* changing from the singular control $b_2/(b_1+b_2)$ to 1. Then the manifold of discontinuity of both U^* and V^* is given by

$$F(t, \mathbf{x}, \mathbf{y}) = t - T + \tau_v(\mathbf{x}, \mathbf{y}) = 0, \tag{29}$$

and

$$G(\mathbf{y}) = a_{11}b_1y_1 - a_{12}b_2y_2 = 0.$$

Across the manifold of discontinuity, we have

$$\mathbf{p}^T(\tau_v^+) = \mathbf{p}^T(\tau_v^-) - \rho \frac{\partial F}{\partial \mathbf{x}} - \sigma \frac{\partial G}{\partial \mathbf{x}},$$

$$\mathbf{q}^T(\tau_v^+) = \mathbf{q}^T(\tau_v^-) - \rho \frac{\partial F}{\partial \mathbf{y}} - \sigma \frac{\partial G}{\partial \mathbf{y}},$$

and

$$H(\tau_v^+) = H(\tau_v^-) + \rho \frac{\partial F}{\partial t} + \sigma \frac{\partial G}{\partial t},$$

or

$$p^T(\tau_v^+) = p^T(\tau_v^-) - \rho \frac{\partial \tau_v}{\partial \mathbf{x}}, \tag{30}$$

$$(-q_1(\tau_v^+)) = (-q_1(\tau_v^-)) + \rho \frac{\partial \tau_v}{\partial y_1} + \sigma a_{11} b_1, \tag{31}$$

$$(-q_2(\tau_v^+)) = (-q_2(\tau_v^-)) + \rho \frac{\partial \tau_v}{\partial y_2} - \sigma a_{12} b_2, \tag{32}$$

and

$$H(\tau_v^+) = H(\tau_v^-) + \rho. \tag{33}$$

If we consider (9) and (11), it is readily shown that

$$S_u(\tau_v^+) = \sigma\{a_{11}(b_1)^2 y_1 + a_{12}(b_2)^2 y_2\} + \rho\left[b_1 y_1 \frac{\partial \tau_v}{\partial y_1} - b_2 y_2 \frac{\partial \tau_v}{\partial y_2}\right], \tag{34}$$

and

$$S_v(\tau_v^+) = -\rho\left[a_{12} x_1 \frac{\partial \tau_v}{\partial x_1} - a_2 \frac{\partial \tau_v}{\partial x_2}\right]. \tag{35}$$

Recalling that $u^*(\tau_v^-) = b_2/(b_1+b_2)$, $u^*(\tau_v^+) = 1$, $v^*(\tau_v^-) = 1$, and $v^*(\tau_v^+) = 0$, we may substitute (30) through (32) into (33) to obtain for $a_{11}x_1y_1 \dfrac{\partial \tau_v}{\partial x_1} + a_2 y_2 \dfrac{\partial \tau_v}{\partial x_2} + b_1 y_1 x_2 \dfrac{\partial \tau_v}{\partial y_1} \neq 1$

$$\rho = \frac{a_{11}(b_1)^2 y_1 x_2 \sigma}{\left[1 - a_{11}x_1y_1 \dfrac{\partial \tau_v}{\partial x_1} - a_2 y_2 \dfrac{\partial \tau_v}{\partial x_2} - b_1 y_1 x_2 \dfrac{\partial \tau_v}{\partial y_1}\right]}. \tag{36}$$

Then we may write

$$S_u(\tau_v^+) = \sigma\left\{a_{11}(b_1)^2 y_1 + a_{12}(b_2)^2 y_2 + \frac{a_{11}(b_1)^2 y_1 x_2\left[b_1 y_1 \dfrac{\partial \tau_v}{\partial y_1} - b_2 y_2 \dfrac{\partial \tau_v}{\partial y_2}\right]}{\left[1 - a_{11}x_1y_1 \dfrac{\partial \tau_v}{\partial x_1} - a_2 y_2 \dfrac{\partial \tau_v}{\partial x_2} - b_1 y_1 x_2 \dfrac{\partial \tau_v}{\partial y_1}\right]}\right\}, \tag{37}$$

and

$$S_v(\tau_v^+) = \frac{-a_{11}(b_1)^2 y_1 x_2 \left[a_{12} x_1 \dfrac{\partial \tau_v}{\partial x_1} - a_2 \dfrac{\partial \tau_v}{\partial x_2} \right] \sigma}{\left[1 - a_{11} x_1 y_1 \dfrac{\partial \tau_v}{\partial x_1} - a_2 y_2 \dfrac{\partial \tau_v}{\partial x_2} - b_1 y_1 x_2 \dfrac{\partial \tau_v}{\partial y_1} \right]}. \tag{38}$$

5. SYNTHESIS OF EXTREMAL STRATEGIC-VARIABLE PAIR

By the synthesis of the extremal strategic-variable pair we mean the explicit determination (using the basic necessary conditions of optimality) of the time history of the extremal strategic-variable pair $(u^*, v^*)^*$ from initial to terminal time [21,31-33]. The basic idea is to trace extremals backwards from the terminal manifold (where boundary conditions for the adjoint variables are known) in such a way as to guarantee the satisfaction of the initial conditions. Thus, it is convenient to introduce the *backwards time* τ defined by (28).

5.1. Extremal Transitions in Strategic Variables

It seems appropriate to examine what the possible transitions (or changes) are in each strategic variable as we work backwards from the end (i.e. as τ increases). It has been shown previously that $\dfrac{\partial S_v}{d\tau} < 0$ for all $\tau \in [0, T]$. If we consider the boundary conditions (4) and (5) for the adjoint variables, it follows that $S_v(\tau = 0) > 0$. Thus

$$v^*(\tau) = \begin{cases} 1 & \text{for } 0 \leqslant \tau \leqslant \tau_v, \\ 0 & \text{for } \tau_v < \tau. \end{cases} \tag{39}$$

It will be convenient to refer to that phase of the planning horizon during which $v^*(t) = 0$ as V-Phase I (i.e. $0 \leqslant t < T - \tau_v$) and to that during which $v^*(t) = 1$ as V-Phase II.

Extremal transitions in u^* for increasing τ are shown in Fig. 3. Thus, this figure shows what changes we might expect to observe in u^* as we follow an extremal backwards from the end of the planning horizon at $\tau = 0$. During V-Phase II when $v^* = 1$, $\dfrac{\partial S_u}{d\tau} = \left(\dfrac{x_1^f}{y_1^f} \right) (a_{11} b_1 y_1 - a_{12} b_2 y_2)$ with $S_u(\tau = 0) = b_1 x_1^f / y_1^f > 0$. When $u^* = 0$, then $\dfrac{d}{d\tau} \left(\dfrac{y_1}{y_2} \right) < 0$. During V-Phase I when $v^* = 0$, $\dfrac{dS_u}{d\tau} = \left(\dfrac{x_1^f}{y_1^f} \right) (a_{11} b_1 y_1 - a_{12} b_2 y_2) + b_2 y_2 S_v(\tau)$. During both phases, the singular subarc may be excited with either $u^* = 0$ or $u^* = 1$. Once u^* becomes 0, it remains this way. The above statements will be further justified below.

5.2. Extremal Synthesis for $\tau_u^* < \tau_v$

From the above we have

$$S_u(\tau = 0) = b_1 \frac{x_1^f}{y_1^f} > 0, \tag{40}$$

*It should be kept in mind that, for example, $u^*(t) = U^*(t, \mathbf{x}, \mathbf{y})$.

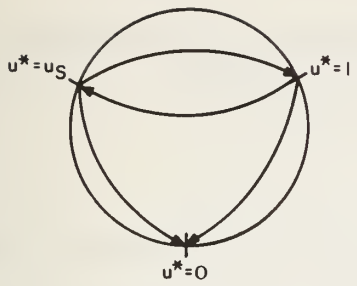


FIGURE 3. Extremal transitions in u* for increasing τ.

NOTE:

$$u_S = \begin{cases} \left[\frac{b_2}{b_1+b_2}\right]\left[1-\frac{q_2 y_2}{p_2 x_2}\right] & \text{FOR } v^*=0, \\ \left[\frac{b_2}{b_1+b_2}\right] & \text{FOR } v^*=1. \end{cases}$$

so that by (8) we have

$$u^*(\tau) = 1 \quad \text{for} \quad 0 \leqslant \tau \leqslant \tau_u, \tag{41}$$

where τ_u is the smallest zero of the equation $S_u(\tau=\tau_u) = 0$. If the U-singular subarc is reached in V-Phase II (see section 4.1. previously), then let us denote the backwards switching time at which u^* changes from 1 to $b_2/(b_1+b_2)$ as τ_u^*. Clearly, it is necessary that $\tau_u^* < \tau_v$ for this singular subarc to appear in the solution. Thus, in general, there are two cases to be considered:

$$\text{(a)} \quad \tau_u^* < \tau_v,$$

and

$$\text{(b)} \quad \tau_u^* \geqslant \tau_v.$$

In this paper we will consider only the former case, with the latter one following along the same general lines of development. We therefore *assume that* $a_{11}, a_{12}, a_2, b_1, b_2, x_1^f, x_2^f, y_1^f$, *and* y_2^f *are such that* $\tau_u^* < \tau_v$. We will give numerical results for this case below. Moreover, in all our numerical computations we have only encountered this case.

5.2.1. Extremals Near the Terminal Manifold

Recalling (16), we see that $\frac{dS_u}{d\tau} > 0$ (<0) if and only if $\frac{y_1}{y_2} > (<) \frac{a_{12}b_2}{a_{11}b_1}$. If we consider (40), it is clear that $S_u(\tau) > 0$ for $v^* = 1$ when $\frac{y_1^f}{y_2^f} > \frac{a_{12}b_2}{a_{11}b_1}$. However, $S_u(\tau)$ may change sign when $\frac{y_1^f}{y_2^f} < \frac{a_{12}b_2}{a_{11}b_1}$. Then U-singular subarc occurs when both $S_u(\tau=\tau_u^*) = 0$ and $a_{11}b_1y_1 = a_{12}b_2y_2$ at $\tau = \tau_u^*$. Thus, τ_u^* is the smallest root of

$$-\frac{1}{b_1 x_2^f} + \tau_u^* + \left(\frac{1}{b_1 x_2^f} - \frac{1}{a_{11} y_1^f}\right) e^{-b_1 x_2^f \tau_u^*} = 0. \tag{42}$$

If y_1^f is given, then $S_u(\tau=\tau_u^*) = 0$ and $a_{11}b_1y_1 = a_{12}b_2y_2$ may be combined to yield the value of y_2^f required in order to reach the U-singular subarc (denoted as y_2^{f*}). Thus, for $a_{11}y_1^f \neq b_2x_2^f$ we have $y_2^f = \frac{b_1}{b_2} \frac{(a_{11}y_1^f - b_1x_2^f)}{a_{12}(1-b_1x_2^f\tau_u^*)}$. (Other results are given in the Appendix.) We denote the corresponding ratio of y_1^f to y_2^{f*} as $\left(\frac{y_1^f}{y_2^f}\right)^*$.

When $S_u(\tau = \tau_u) = 0$ with $a_{11}b_1y_1 < a_{12}b_2y_2$, it follows that τ_u is the smallest root of the transcendental equation

$$\left(b_1 - \frac{a_{11}y_1^f}{x_2^f}\right) - a_{12}b_2y_2^f\tau_u + \frac{a_{11}y_1^f}{x_2^f}\, e^{b_1x_2^f\tau_u} = 0. \tag{43}$$

THEOREM 1: Assume that $\tau_v > \tau_u^*$. Then, $u^*(\tau) = 1$ on any extremal as long as $v^*(\tau) = 1$ for $\frac{y_1^f}{y_2^f} > \left(\frac{y_1^f}{y_2^f}\right)^*$.

PROOF: The proof is by contradiction. Let $r = y_1^f/y_2^f$.

(a) Assume that we could have a switch in $u^*(\tau)$ (with $v^*(\tau) = 1$) for $\frac{y_1^f}{y_2^f} > \left(\frac{y_1^f}{y_2^f}\right)^*$. In other words, we can find τ_u such that $S_u(\tau{=}\tau_u) = 0$ with

$$S_u(\tau) > 0 \text{for } 0 \leqslant \tau < \tau_u, \tag{44}$$

for $\frac{y_1^f}{y_2^f} > \left(\frac{y_1^f}{y_2^f}\right)^*$.

(b) Consider $\tilde{r} = \frac{y_1^f}{y_2^f} = \left(\frac{y_1^f}{y_2^f}\right)^* + \epsilon$ with $\epsilon > 0$ and such that $\tau_u < \tau_v$. Then it may be shown that at $r = (y_1^f/y_2^f)^*$ we have $\partial\tau_u/\partial r > 0$. This implies, however, that $\tau_u > \tau_u^*$ for $r = \tilde{r}$.

(c) Observe that $\begin{cases} u^* = 1 \\ v^* = 1 \end{cases}$ for $0 \leqslant \tau \leqslant \tau_u < \tau_v$ so that $y_1/y_2 = y_1^f/y_2^f e^{b_1x_2^f\tau}$. Hence,

$$\frac{y_1}{y_2}\,(\tau{=}\tau_u) > \frac{a_{12}b_2}{a_{11}b_1} \text{ for } r = \tilde{r}, \tag{45}$$

since then $\tau_u(r{=}\tilde{r}) > \tau_u^*$. It has been shown above that $\frac{dS_u}{d\tau} > 0$ for $y_1/y_2 > a_{12}b_2/(a_{11}b_1)$. Thus, (45) implies that $\frac{dS_u}{d\tau}\,(\tau{=}\tau_u) > 0$, and hence

$$0 = S_u(\tau{=}\tau_u) > S_u(\tau) \text{ for } \tau \in (\tau_u - \delta, \tau_u). \tag{46}$$

This last statement (46) is a contradiction to (44), and the theorem is proved

Q.E.D.

Other results are obtained in a similar fashion.

5.2.2. Field Construction

For a given set of terminal values x_1^f, x_2^f, y_1^f and y_2^f, an extremal may be traced backwards from the terminal manifold by a backwards integration of the state and adjoint equations combined with (8) and (10) (also (20) or (26)). By varying these terminal values, the entire *field of extremals* (see p. 128 of Ref. [12]) may be obtained.

The various types of extremals that may occur in the field of extremals are shown in Fig. 4. This figure is representative of all our numerical results for $\tau_u^* < \tau_v$ (see Section 5.2.3 below). Pertinent information concerning each type of extremal is given in the Appendix.

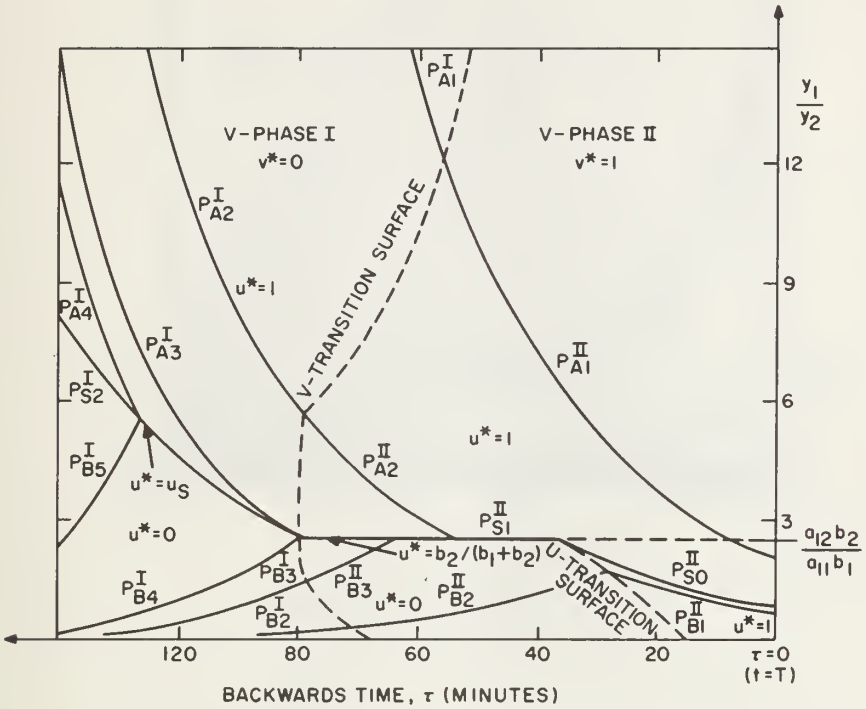


FIGURE 4. Identification of various types of extremals for which information is given in the Appendix.

5.2.3. Numerical Examples

A computer program to calculate numerical values for information given in the Appendix was written in FORTRAN for the IBM 360 computer.* A plot of the field of extremals (see Fig. 5,6, and 7 below) is generated by this program. The closed-form analytic results presented in the Appendix are used whenever possible. Approximate numerical solutions to transcendental equations (for the determination of, for example, τ_u^*, τ_v, etc.) are developed by the well-known Newton-Raphson method. In those cases for which closed-form solutions are not available to the state and adjoint equations, a standard fourth-order Runge-Kutta numerical integration method is used. A time step $\Delta\tau$ was used in these numerical integrations which yielded agreement to the fifth place to the right of the decimal place in tests cases in which the approximate numerical solution could be compared with the exact solution.

Parameter sets for the numerical examples given in this paper are shown in Table 1. for our problem (2) we may consider time to be an additional state variable so that the state space is five-dimensional, i.e., the state variables are t, x_1, x_2, y_1, and y_2. Thus, unfortunately, we cannot graphically depict the field of extremal trajectories but must be satisfied with viewing "cross-section" plots of it.

*The author would like to thank Captain Jeffrey L. Ellis (U.S. Army) for doing this work. Subsequent computational contributions were made by Captain Robert J. Hill III (U.S. Army).

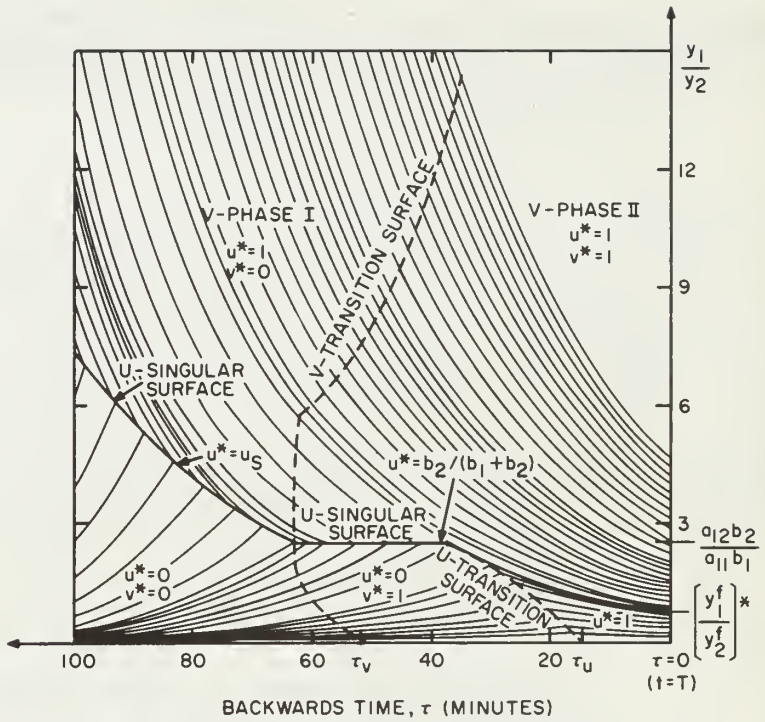


FIGURE 5. Plot of y_1/y_2 vs backwards time σ for field of extremals for parameter set 1.

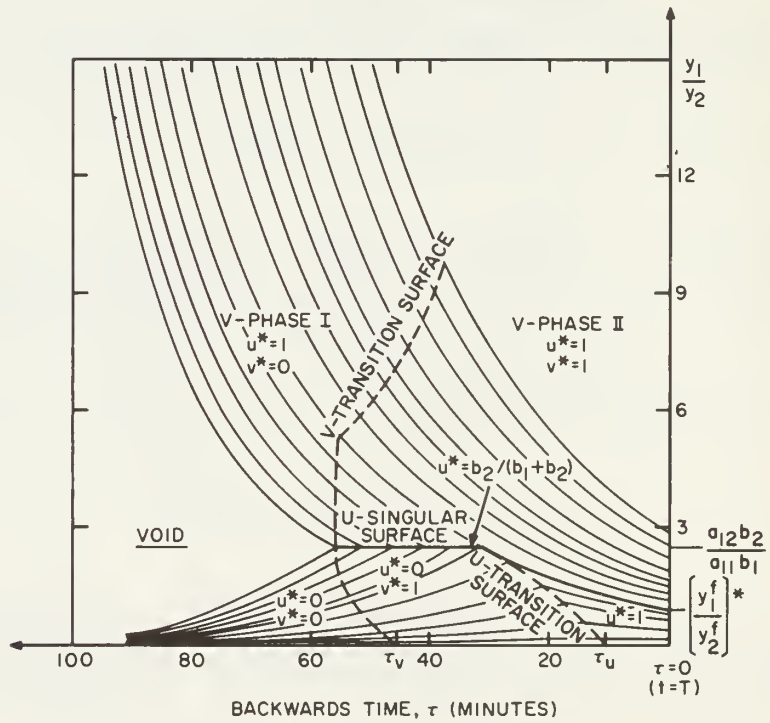


FIGURE 6. Void in field of extremals shown in plot of y_1/y_2 vs backwards time τ for parameter set 2.

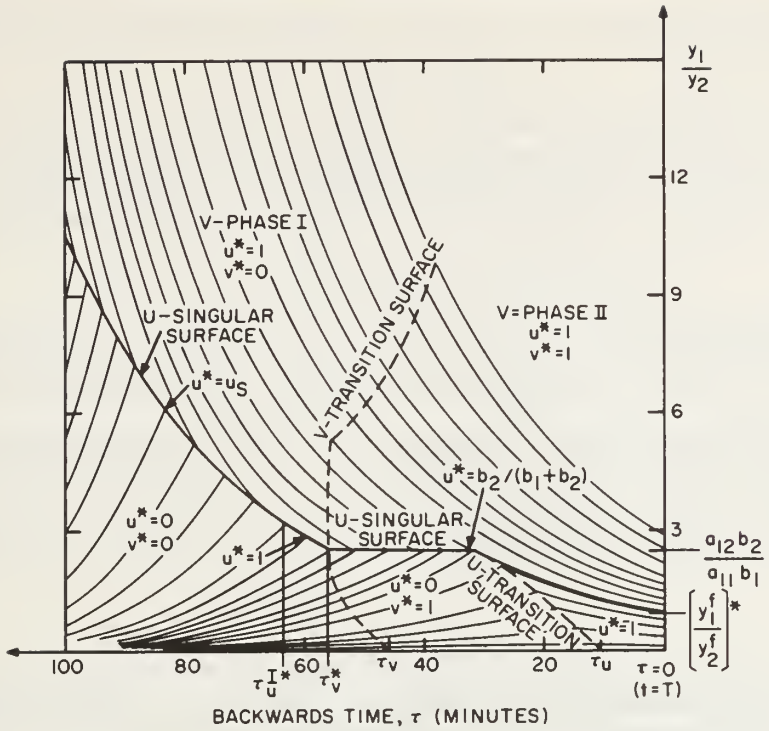


FIGURE 7. Filled-in void in field of extremals for parameter set 2.

TABLE 1. *Parameter Sets Used to Generate Numerical Results Shown in figs. 5,6, and 7.*

Parameter Set	a_{11}	a_{12}	a_2	b_1	b_2	x_1^f	x_2^f	y_2^f
1	0.003	0.006	0.01	0.004	0.005	4.0	8.0	8.964
2	0.003	0.006	0.01	0.004	0.005	4.0	8.0	11.597

The most illuminating plot for gaining insight into the structure of the optimal fire-support strategies for (2) is that of extremal trajectories in terms of y_1/y_2 vs backwards time τ. This is shown for parameter set 1 in Fig. 5. The corresponding strategic variable values for X and Y (i.e. u^* and v^*) along each extremal are also given. Other plots have been considered, but they provide little, if any, additional insight.

The most significant features of the field of extremals shown in Fig. 5 are the two U-singular "surfaces"; there is one in $\mathbf{x}, \mathbf{y} - \mathbf{p}, \mathbf{q}$ space in V-phase I and one in $\mathbf{y}$-space in V-phase II. In each phase, X uses the strategy $U^* = 1$ above the singular "surface" and the strategy $U^* = 0$ below it. Similar to our discussion in Ref. [32], the singular surfaces are present in the field of optimal trajectories so that the X artillery avoids "overkilling" either Y_1 or Y_2. This insight is obvious when one, for example, considers

$$\frac{\left(-\frac{dy_1}{dt}\right)}{x_2} = b_1 y_1.$$

Thus, the rate of destruction of Y_1 per unit of X artillery decreases over time as the Y_1 force level decreases [31,32].

Results for parameter set 2 are shown in Fig. 6. There is a void (see p. 169 and also p. 187 of Ref. [21]) in the field of extremals. This is because in backwards time at the end τ_v^* of the U-singular subarc in V-phase II, we would have $u_s(\tau_v^{*+})$ (as given in Fig. 3) equal to 1.054 if the adjoint variables were continuous at τ_v^*. The following theorem further explains this situation.

THEOREM 2: There can be no U-singular subarc beginning in backwards time at τ_v^+ with $a_{11}b_1y_1 = a_{12}b_2y_2$ for

$$b_1p_2(\tau_v^-)x_2 + b_2q_2(\tau_v^-)y_2 < 0.$$

When a U-singular subarc begins at τ_v^+ with $a_{11}b_1y_1 = a_{12}b_2y_2$, there is no discontinuity in the adjoint variables at $\tau = \tau_v$ (i.e. $\sigma = 0$ in (37)).

PROOF: Immediate by (27) and (37).

Q.E.D.

Additionally, Theorem 3 gives the extremal transitions in X's strategy possible from the U-singular surface in V-phase II as we work backwards from τ_v^*. Thus, since $b_1p_2(\tau_v^*)x_2 < b_2(-q_2(\tau_v^*))y_2$ for parameter set 2, a void would exist in the field of extremals if the adjoint variables were continuous at τ_v^*.

THEOREM 3: Assume that there is no discontinuity in the adjoint variables at $\tau = \tau_v$ with $a_{11}b_1y_1 = a_{12}b_2y_2$. Then

I. if $b_1p_2(\tau_v)x_2 < b_2(-q_2(\tau_v))y_2$, then we can only have $u^*(\tau) = 0$ for $\tau \in (\tau_v, \tau_v+\delta)$ where $\delta > 0$,

II. If $b_1p_2(\tau_v)x_2 \geqslant b_2(-q_2(\tau_v))y_2$, then we can have

$$u^*(\tau) = \begin{cases} (a)\ 0, \\ (b)\ (1 - q_2y_2/(p_2x_2))\cdot b_2/(b_1+b_2), \\ (c)\ 1, \end{cases}$$

for $\tau \in (\tau_v, \tau_v+\delta)$ where $\delta > 0$.

PROOF: (a) When we are on the singular surface in V-phase II at $\tau_v^- = \tau_v^{*-}$, then by (22) and (23) and the continuity of the dual variables we have

$$S_u(\tau = \tau_v^+) = \mathring{S}_u(\tau = \tau_v^+) = 0,$$

and

$$\mathring{\mathring{S}}_u(\tau = \tau_v^+) = a_2b_2(b_1+b_2)p_2x_2y_2\left\{u^*(\tau_v^+) - \left(\frac{b_2}{b_1+b_2}\right)\left(1 - \frac{q_2y_2}{p_2x_2}\right)\right\}, \tag{47}$$

where $\mathring{S}$ denotes $\dfrac{dS_u}{d\tau}$.

(b) Considering a Taylor series expansion about $\tau = \tau_v^+$, we have by the above for $\tau \geqslant \tau_v^+ = \tau_v^{*+}$

$$S_u(\tau) = \frac{(\tau - \tau_v^*)^2}{2} \overset{\circ\circ}{S}_u(\bar{\tau}), \tag{48}$$

where $\bar{\tau} \in (\tau_v^*, \tau)$.

(c) When $u^*(\tau) = 0$ for $\tau \in (\tau_v^*, \tau_v^* + \delta)$, then

$$\overset{\circ\circ}{S}_u(\tau = \tau_v^{*+}) = -a_2(b_2)^2 p_2 x_2 y_2 \left[1 - \frac{q_2 y_2}{p_2 x_2}\right] < 0,$$

so that there exists $\delta_1 > 0$ such that $S_u(\tau) < 0$ for all $\tau \in (\tau_v^*, \tau_v^* + \delta_1)$. Thus, we can always have $u^* = 0$ as we work backwards in V-phase I from the U-singular subarc in V-phase II.

(d) Now let $b_1 p_2(\tau_v^*) x_2 \geqslant b_2(-q_2(\tau_v^*)) y_2$. By (26), the U-singular control in V-phase I $u_S = (1 - q_2 y_2/(p_2 x_2)) \cdot b_2/(b_1 + b_2) \leqslant 1$. Thus, the U-singular subarc is possible. When $u^*(\tau_v^{*+}) = 1$, then $\overset{\circ\circ}{S}_u(\tau = \tau_v^{*+}) \geqslant 0$ by (47). When inequality holds, it follows that there exists $\delta_1 > 0$ such that $S_u(\tau) > 0$ for all $\tau \in (\tau_v^*, \tau_v^* + \delta_1)$. Clearly, we cannot have $u^* = 1$ if $b_1 p_2(\tau_v^*) x_2 < b_2(-q_2(\tau_v^*)) y_2$.

Q.E.D.

The same analysis as used in the proof of Theorem 3 applies on a U-singular subarc in V-phase I when $v^* = 0$. As long as (27) holds, one has three options similar to those of part II of Theorem 3.

5.2.4. Filling in a Void

We have emphasized that H, $\mathbf{p}(t)$, and $\mathbf{q}(t)$ are continuous functions of time except possibly at manifolds of discontinuity of both U^* and V^* (see Section 4.3). From Theorem 3 it follows that a void must exist in the field of extremals when these functions are continuous and $b_1 p_2(\tau_v^*) x_2 < b_2(-q_2(\tau_v^*)) y_2$. At τ_v^*, moreover, v^* changes (as we progress backwards in time) from 1 to 0 and u^* from $b_2/(b_1 + b_2)$ to a different value. Thus, we have a manifold of discontinuity of both U^* and V^*. Moreover, if we consider results given above, it is readily shown that $u^*(\tau)$ remains for increasing τ (i.e., backwards time) equal to zero once it changes to zero. Then from Theorems 2 and 3 it follows that for $b_1 p_2(\tau_v^{*-}) x_2 < b_1(-q_2(\tau_v^{*-})) y_2$, the dual variables must be discontinuous to fill in the void, and we must have $u^*(\tau) = 1$ for $\tau_v^* < \tau < \tau_u^{I*}$. Furthermore, if we consider Fig. 6 and considerations "in the large," the manifold of discontinuity must lie on the V-transition surface.

Thus, we have established that for $a_{11} b_1 y_1 = a_{12} b_2 y_2$ we have

$$\begin{cases} u^*(\tau_v^-) = b_2/(b_1 + b_2), \\ v^*(\tau_v^-) = 1 \end{cases}$$

and

$$\begin{cases} u^*(\tau_v^+) = 1, \\ v^*(\tau_v^+) = 0. \end{cases} \tag{49}$$

It remains to determine the function $\tau_v(\mathbf{x}, \mathbf{y})$ of (29) so that $\dfrac{\partial \tau_v}{\partial \mathbf{x}}$ and $\dfrac{\partial \tau_v}{\partial \mathbf{y}}$ may be computed, and the jumps in H, $\mathbf{p}$, and $\mathbf{q}$ subsequently determined (see (30) through (33)). It should be clear that it is impossible to explicitly determine $\tau_v(\mathbf{x}, \mathbf{y})$. However, by computation of five points on the V-transition surface, the desired partial derivatives may be estimated by using

linear approximations to the appropriate directional derivatives and solving a system of four linear equations in four unknowns. For parameter set 2 (as the reference case), this yielded the following estimates

$$\frac{\partial \tau_v}{\partial x_1} = 0.0000, \qquad \frac{\partial \tau_v}{\partial x_2} = -0.295,$$

(50)

$$\frac{\partial \tau_v}{\partial y_1} = -0.0167, \qquad \frac{\partial \tau_v}{\partial y_2} = -0.0331.$$

It is, therefore, convenient to rewrite the jump conditions across the manifold of discontinuity of both U^* and V^*.

$$p_1(\tau_v^+) = p_2(\tau_v^-), \qquad p_1(\tau_v^+) = p_2(\tau_v^-) - \rho \frac{\partial \tau_v}{\partial x_2},$$

(51)

$$q_1(\tau_v^+) = q_1(\tau_v^-) - \rho \frac{\partial \tau_v}{\partial y_1} - \sigma a_{11} b_1,$$

$$q_2(\tau_v^+) = q_2(\tau_v^-) - \rho \frac{\partial \tau_v}{\partial y_2} + \sigma a_{12} b_2,$$

where ρ and σ are related by (36). In this case the jumps (37) and (38) in the switching functions simplify to

$$S_u(\tau_v^+) = \sigma \left\{ a_{11}(b_1)^2 y_1 + a_{12}(b_2)^2 y_2 + \frac{a_{11}(b_1)^2 y_1 x_2 \left(b_1 y_1 \dfrac{\partial \tau_v}{\partial y_1} - b_2 y_2 \dfrac{\partial \tau_v}{\partial y_2} \right)}{\left(1 - a_2 y_2 \dfrac{\partial \tau_v}{\partial x_2} - b_1 y_1 x_2 \dfrac{\partial \tau_v}{\partial y_1} \right)} \right\}. \tag{52}$$

and

$$S_v(\tau_v^+) = \frac{a_{11} a_2 (b_1)^2 y_1 x_2 \dfrac{\partial \tau_v}{\partial x_2} \sigma}{\left(1 - a_2 y_2 \dfrac{\partial \tau_v}{\partial x_2} - b_1 y_1 x_2 \dfrac{\partial \tau_v}{\partial y_1} \right)}. \tag{53}$$

Since $v^*(\tau_v^{*+}) = 0$, we must have $S_v(\tau_v^{*+}) \leqslant 0$ so that (50) and (53) yield that $\sigma \geqslant 0$. It should be clear that $\sigma = 0$ if and only if **H**, **p**, and **q** are continuous at τ_v^*. For $\sigma > 0$, the condition that $u^*(\tau_v^{*+}) = 1$ yields that we must have

$$\frac{S_u(\tau_v^{*+})}{\sigma} > 0, \tag{54}$$

where $S_u(\tau_v^{*+})$ is given by (52). Although it cannot in general be guaranteed that (54) will always hold when a void in the field of extremals such as that shown in Fig. 6 exists, it should be clear that it must if the problem (2) is to have a solution. The author conjectures that this is true. It is readily shown that when (54) holds, we have

$$S_u(\tau_v^{*+}) > 0, \ \mathring{S}_u(\tau_v^{*+}) < 0, \text{ and } S_v(\tau_v^{*+}) < 0. \tag{55}$$

The appropriate value for σ is determined by "considerations in the large": the structure of the entire field of extremals determines the value of this parameter. In Fig. 7, we let $\tau_u^{l^*}$

denote the backwards time at which the U-singular subarc is entered in V-phase I. Corresponding to $\tau_v^{I^*}$ is σ^*, which yields the first and second conditions (18) and (25) (with $u_S^* \leqslant 1$) for a U-singular subarc with $V^* = 0$ at $\tau_u^{I^*} > \tau_v^*$. For $0 < \sigma < \sigma^*$, one uses $u^*(\tau) = 1$ for $\tau_v^{*+} < \tau < \tau_u^I$ and then $u^*(\tau) = 0$ for $\tau > \tau_u^I$. For $\sigma > \sigma^*$, the U-switching function $S_u(\tau)$ never changes sign, so that $u^*(\tau) = 1$ for all $\tau > \tau_v^*$. Thus, by manipulation of σ, one may fill in the void in the field of extremals in V-Phase I. The resulting field of extremals is shown in Fig. 7.

5.2.5. The Case of Negligible Y_1 Small-Arms Effectiveness

It seems appropriate to consider what happens to the solution to the problem at hand as the (relative) effectiveness of Y_1 (small-arms) fire becomes negligible, i.e. as $a_{11} \to 0$. Let us consider Fig. 5 (or Fig. 7). The U-singular "surface" in V-Phase II has equation $y_1/y_2 = a_{12}b_2/(a_{11}b_1)$. Thus, as $a_{11} \to 0$ with the other parameters being held constant, this singular "surface" appears higher and higher on the y_1/y_2 axis in Fig. 5. In the limit, the singular surface does not appear in the finite part of the plane. Thus, we have shown that an optimal strategy in which a side divides the fire of its supporting weapon system between the enemy's primary (infantry) and supporting systems *can only occur* when the enemy's infantry has some fire effectiveness (in the sense of a nonzero Lanchester attrition-rate coefficient) against his infantry.

6. DISCUSSION

In this paper we have examined the dependence of optimal time-sequential fire-support strategies on the form of the combat attrition model by considering a differential game (see (2)) with slightly different combat dynamics than those in the fire-support differential game considered by Kawara [22] (see (1)). For this fire-support differential game (2) we developed first-order necessary conditions of optimality and constructed "cross-section" pictures of the field of extremals. By comparing and contrasting the structure of optimal fire-support strategies for our problem (2) with that for Kawara's fire-support differential game (1), one begins to understand the nature of the dependence of optimal strategies on the combat dynamics by also comparing and contrasting the combat attrition equations for these two differential games.

Our fire-support differential game (2) was similar to Kawara's problem (1) (see Ref. [22]) except that we let the attacker's (i.e. X's) artillery produce "linear-law" attrition* against both the defender's artillery and also his infantry and let the defender's infantry produce "linear-law" attrition against the attacker's infantry. As contrasted with the optimal time-sequential fire-support strategies for Kawara's problem (1) of always concentrating all artillery fire first on enemy artillery and later on enemy infantry (the timing of the switch being force-level independent), for our problem (2) the optimal strategy for one combatant (the attacker, X) depends directly on the enemy's force levels and is no longer to always concentrate all fire on either the enemy's primary or secondary weapon system. The latter result, moreover, was shown to depend on the defender's infantry having some fire effectiveness (in the sense of a non-zero Lanchester attrition-rate coefficient) against the attacker's infantry.

The solution to (2) is characterized by the presence of singular surfaces (in Issacs' terminology [21], universal surfaces [18], a different one for each V-phase of battle. When the battle state reaches one of these surfaces, X follows an optimal strategy of dividing his artillery fire between enemy infantry and artillery in order to avoid "overkill." Another characteristic of the optimal fire-support strategies (not present for Kawara's [22] problem (1)) is that X's optimal

*For convenience we use the term "linear-law" attrition to denote an attrition process in which a target type undergoes attrition at a rate proportional to the product of the numbers of firers and targets [31,32].

strategy may sometimes depend on Y's distribution of supporting fires. This behavior occurs on the singular surfaces. In fact, X sometimes must react instantaneously to changes in Y's fire distribution.

The development of even a partial solution to (2) has involved a solution phenomenon not previously reported for Lanchester-type differential games: the adjoint (or dual) variables* are discontinuous across a manifold of discontinuity of both U^* and V^*. This manifold of discontinuity exists for a certain range of parameter values in the solution to the problem at hand (2). Furthermore, there is a military interpretation to this manifold of discontinuity: if Y_2 concentrates fire on X_2 and X_2 on Y_1, then when Y_2 changes to concentrating all fire on X_1, X must re-evaluate the worth of a Y_2 unit because it now has a direct influence on the payoff. Such a discontinuity in the adjoint variables is unique to differential games [3,4] (i.e., it cannot occur for a one-sided optimal control problem).

It should also be pointed out that the presence of singular (i.e. universal) surfaces in the solution to (2) is apparently independent of the form of the criterion functional (here, terminal payoff) and depends only on the combat dynamics. For purposes of comparison we considered the same payoff as considered by Kawara [22]. We also showed that the singular (i.e., universal) surfaces can only by present in the solution when the defender's infantry Y_1 has a nonzero casualty producing capability against X_1.

The problem (2) considered in this paper has certain similarities to the "War of Attrition and Attack: Second Version" studied by R. Isaacs (see pp. 330-335 of Ref. [21]). We have, however, developed a much more complete solution to our problem than that given in Ref. [21] for Mengel's problem. Although this problem (2) possesses some similarities to the Lanchester-type optimal control problem studied by us in Ref. [31], its solution has turned out to be much more complex. Our developments in this paper, however, have been significantly helped by intuition gained in the study of the simpler, one-sided problem (see Ref. [32] for a further discussion).

As a result of our investigation here, we hope that a better understanding of optimal fire-support strategies has been developed. As is always the case, however, the insights gained into the optimization of combat dynamics from our study of the differential game (2) are no more valid than the combat model itself. Our work here shows that the functional forms of the various target-type casualty rates produced by the artillery essentially determine the most significant aspects of the structure of the optimal fire-support strategies. Thus, our study of this optimization problem shows the importance of determining the appropriate (Lanchester-type) model of combat dynamics.

APPENDIX

Extremal Trajectories for the Fire-Support Problem with $\tau_u^* < \tau_v$:

In this appendix we give information about the various types of extremals shown in Fig. 4.

*The reader should recall that these represent the marginal values of force types, i.e., $p_2(t) = \partial V/\partial x_2(t)$ where $V = V(t,\mathbf{x},\mathbf{y})$ denotes the value of the differential game (see [14,21]).

E1.
$$P_{S0}^{II}: \begin{cases} u^*(\tau) = 1 \\ \qquad\qquad \text{for } 0 \leqslant \tau \leqslant \tau_u^* \text{ with } \dfrac{y_1^f}{y_2^f} = \left[\dfrac{y_1^f}{y_2^f}\right]^* \\ v^*(\tau) = 1 \end{cases}.$$

τ_u^* is the smallest positive root of

$$-\frac{1}{b_1 x_2^f} + \tau_u^* + \left[\frac{1}{b_1 x_2^f} - \frac{1}{a_{11} y_1^f}\right] e^{-b_1 x_2^f \tau_u^*} = 0,$$

with the following bounds established:

$$\text{for } a_{11}y_1^f > b_1 x_2^f\text{: } \frac{1}{a_{11}y_1^f} < \tau_u^* < \frac{1}{b_1 x_2^f},$$

$$\text{for } a_{11}y_1^f = b_1 x_2^f\text{: } \tau_u^* = \frac{1}{b_1 x_2^f},$$

$$\text{for } a_{11}y_1^f < b_1 x_2^f\text{: } \frac{1}{b_1 x_2^f} < \tau_u^* < \frac{1}{a_{11}y_1^f};$$

$$\text{for } a_{11}y_1^f \neq b_1 x_2^f\text{: } y_2^{f*} = \frac{b_1}{b_2}\,\frac{\left(a_{11}y_1^f - b_1 x_2^f\right)}{a_{12}\left(1 - b_1 x_2^f \tau_u^*\right)},$$

$$\text{for } a_{11}y_1^f = b_1 x_2^f\text{: } y_2^{f*} = \frac{a_{11}b_1 y_1^f}{a_{12}b_2}\, e.$$

Also,

$$S_v(\tau) = a_2 b_1 \left[\frac{x_1^f}{y_1^f}\right]\left\{\left[\left[\frac{a_{12}}{a_2 b_1} + \frac{a_{11}y_1^f}{(b_1 x_2^f)^2}\right] + \left[\frac{a_{11}y_1^f}{b_1 x_2^f} - 1\right]\tau - \frac{a_{11}y_1^f}{(b_1 x_2^f)^2}\, e^{b_1 x_2^f \tau}\right\}.$$

Let $S_v(\tau{=}\tau_u^*) = S_v^{u*}$. Also, on P_{S0}^{II} we have

$$x_1(\tau) = x_1^f \exp\left\{a_{12}y_2^f\tau + \frac{a_{11}y_1^f}{b_1 x_2^f}\left[e^{b_1 x_2^f \tau} - 1\right]\right\},$$

$$x_2(\tau) = x_2^f,$$

$$y_1(\tau) = y_1^f\, e^{b_1 x_2^f \tau},$$

$$y_2(\tau) = y_2^f,$$

and

$$p_1(\tau) = \frac{1}{y_1^f} \exp\left\{-\left[a_{12}y_2^f\tau + \frac{a_{11}y_1^f}{b_1 x_2^f}\left[e^{b_1 x_2^f \tau} - 1\right]\right]\right\},$$

$$p_2(\tau) = b_1\left[\frac{x_1^f}{y_1^f}\right]\left\{\left[1 - \frac{a_{11}y_1^f}{b_1 x_2^f}\right]\tau + \frac{a_{11}y_1^f}{(b_1 x_2^f)^2}\left[e^{b_1 x_2^f \tau} - 1\right]\right\},$$

$$q_1(\tau) = \left[\frac{x_1^f}{y_1^f}\right]\left\{-\frac{a_{11}}{b_1 x_2^f} + \left[\frac{a_{11}}{b_1 x_2^f} - \frac{1}{y_1^f}\right] e^{-b_1 x_2^f \tau}\right\},$$

$$q_2(\tau) = -a_{12}\left[\frac{x_1^f}{y_1^f}\right]\tau.$$

E2. $$P_{S1}^{II}: \begin{cases} u^*(\tau) = b_2/(b_1+b_2) \\ \qquad\qquad \text{for } \tau_u^* \leqslant \tau \leqslant \tau_v^*, \text{ where } \tau_u^* \text{ is determined in E1.} \\ v^*(\tau) = 1 \end{cases}$$

On P_{S1}^{II} we have

$$S_u(\tau) = 0,$$

and

$$a_{11}b_1y_1 = a_{12}b_2y_2.$$

τ_v^* is the smallest positive root of $S_v(\tau=\tau_v^*) = 0$, where

$$S_v(\tau) = S_v^{u^*} + a_2b_1y_1^{u^*}\left(\frac{x_1^f}{y_1^f}\right)\left\{\frac{a_{11}}{(\theta x_2^f)^2} + \left[q_1^{u^*}\left(\frac{y_1^f}{x_1^f}\right) + \frac{a_{11}}{\theta x_2^f}\right](\tau - \tau_u^*) - \frac{a_{11}}{(\theta x_2^f)^2}\exp\left[\theta x_2^f(\tau - \tau_u^*)\right]\right\},$$

with $\theta = b_1b_2/(b_1+b_2)$. An upper bound on τ_v^* is given by

$$\hat{\tau}_v^* = a_{12}/(a_2b_1).$$

Also, on P_{S1}^{II} we have

$$x_1(\tau) = x_1^{u^*}\exp\left\{\left[\frac{a_{11}y_1^{u^*} + a_{12}y_2^f}{\theta x_2^f}\right]\left[e^{\theta x_2^f(\tau-\tau_u^*)} - 1\right]\right\},$$

$$x_2(\tau) = x_2^f,$$

$$y_1(\tau) = y_1^{u^*}\, e^{\theta x_2^f(\tau-\tau_u^*)},$$

$$y_2(\tau) = y_2^f\, e^{\theta x_2^f(\tau-\tau_u^*)},$$

and

$$p_1(\tau) = p_1^{u^*}\exp\left\{-\left[\frac{a_{11}y_1^{u^*} + a_{12}y_2^f}{\theta x_2^f}\right]\left[e^{\theta x_2^f(\tau-\tau_u^*)} - 1\right]\right\} \text{ with } p_1^{u^*} = \frac{1}{x_1^{u^*}}\left(\frac{x_1^f}{y_1^f}\right),$$

$$p_2(\tau) = p_2^{u^*} - b_1y_1^{u^*}\left(\frac{x_1^f}{y_1^f}\right)\left\{\frac{a_{11}}{(\theta x_2^f)^2} + \left[q_1^{u^*}\left(\frac{y_1^f}{x_1^f}\right) + \frac{a_{11}}{\theta x_2^f}\right](\tau-\tau_u^*) - \frac{a_{11}}{(\theta x_2^f)^2}\, e^{\theta x_2^f(\tau-\tau_u^*)}\right\},$$

$$q_1(\tau) = q_1^{u^*}\, e^{-\theta x_2^f(\tau-\tau_u^*)} - \left(\frac{x_1^f}{y_1^f}\right)\frac{a_{11}}{\theta x_2^f}\left\{1 - e^{-\theta x_2^f(\tau-\tau_u^*)}\right\}.$$

$$q_2(\tau) = q_2^{u^*}\, e^{-\theta x_2^f(\tau-\tau_u^*)} - \left(\frac{x_1^f}{y_1^f}\right)\frac{a_{12}}{\theta x_2^f}\left\{1 - e^{-\theta x_2^f(\tau-\tau_u^*)}\right\}.$$

E3.
$$P_{A1}^{II}: \begin{cases} u^*(\tau) = 1 \\ \qquad \text{for } 0 \leqslant \tau \leqslant \tau_v \text{ with } \dfrac{y_1^f}{y_2^f} > \left(\dfrac{y_1^f}{y_2^f}\right)^* \\ v^*(\tau) = 1 \end{cases}.$$

τ_v is the smallest positive root of $S_v(\tau{=}\tau_v) = 0$, where $S_v(\tau)$ is given in E1. An upper bound on τ_v is given by

$$\hat{\tau}_v = \frac{a_{12}}{a_2 b_1}.$$

It has been shown that $S_u(\tau) > 0$ for $0 \leqslant \tau \leqslant \tau_v$. The solutions to the state and adjoint equations are the same as those for P_{S0}^{II} given previously.

E4.
$$P_{A1}^{I}: \begin{cases} u^*(\tau) = 1 \\ \qquad \text{for } \tau_v \leqslant \tau \leqslant \tau_u^I. \\ v^*(\tau) = 0 \end{cases}$$

We have that $S_v(\tau) < 0$ for $\tau > \tau_v$ and that

$$\frac{dS_u}{d\tau}(\tau) = b_2 y_2^f S_v(\tau) + \left(\frac{x_1^f}{y_1^f}\right)(a_{11} b_1 y_1(\tau) - a_{12} b_2 y_2^f).$$

Also, on P_{A1}^I we have

$$\frac{dx_1}{d\tau} = a_{11} x_1 y_1 \text{ with } x_1(\tau{=}\tau_v) = x_1^v,$$

$$x_2(\tau) = x_2^v + a_2 y_2^f(\tau{-}\tau_v),$$

$$y_1(\tau) = y_1^v \exp\left\{b_1 x_2^v(\tau{-}\tau_v) + \frac{a_2 b_1 y_2^f}{2}(\tau{-}\tau_v)^2\right\},$$

$$y_2(\tau) = y_2^f,$$

and

$$\frac{dp_1}{d\tau} = -a_{11} y_1 p_1 \text{ with } p_1(\tau{=}\tau_v) = p_1^v,$$

$$\frac{dp_2}{d\tau} = -b_1 y_1 q_1 \text{ with } p_2(\tau{=}\tau_v) = p_2^v,$$

$$\frac{dq_1}{d\tau} = -a_{11}\frac{x_1^f}{y_1^f} - b_1 x_2 q_1 \text{ with } q_1(\tau = \tau_v) = q_2^v,$$

$$\frac{dq_2}{d\tau} = -a_2 p_2 \text{ with } q_2(\tau{=}\tau_v) = q_2^v.$$

We have not been able to develop solutions in terms of "elementary" functions to the equations for x_1, p_1, p_2, q_1, and q_2.

E5. $$P^{II}_{A2}: \begin{cases} u^*(\tau) = 1 \\ \qquad\qquad \text{for } \tau^{II}_{SL} \leqslant \tau \leqslant \tau_v. \\ v^*(\tau) = 1 \end{cases}$$

τ_v is the smallest positive root of $S_v(\tau{=}\tau_v) = 0$, where

$$S_v(\tau) = S_v^{SLII} + a_2 b_1 y_1^{SLII} \left[\frac{x_1^f}{y_1^f}\right] \left\{\frac{a_{11}}{(b_1 x_2^f)^2} + \left[q_1^{SLII}\left[\frac{y_1^f}{x_1^f}\right] + \frac{a_{11}}{b_1 x_2^f}\right](\tau - \tau^{II}_{SL}) - \frac{a_{11}}{(b_1 x_2^f)^2} e^{b_1 x_2^f (\tau - \tau^{II}_{SL})}\right\}.$$

Again, an upper bound on τ_v is given by $a_{12}/(a_2 b_1)$. It has been shown that $S_u(\tau) > 0$ for $\tau^{II}_{SL} < \tau \leqslant \tau_v$. Also, on P^{II}_{A2} we have

$$x_1(\tau) = x_1^{SLII} \exp\left\{a_{12} y_2^{SLII}(\tau - \tau^{II}_{SL}) + \frac{a_{11} y_1^{SLII}}{b_1 x_2^f}\left[e^{b_1 x_2^f (\tau - \tau^{II}_{SL})} - 1\right]\right\},$$

$$x_2(\tau) = x_2^f,$$

$$y_1(\tau) = y_1^{SLII} e^{b_1 x_2^f (\tau - \tau^{II}_{SL})},$$

$$y_2(\tau) = y_2^{SLII},$$

and

$$p_1(\tau) = p_1^{SLII} \exp\left\{-a_{12} y_2^{SLII}(\tau - \tau^{II}_{SL}) - \frac{a_{11} y_1^{SLII}}{b_1 x_2^f}\left[e^{b_1 x_2^f (\tau - \tau^{II}_{SL})} - 1\right]\right\}$$

$$\text{with } p_1^{SLII} = \frac{1}{x_1^{SLII}}\left[\frac{x_1^f}{y_1^f}\right],$$

$$p_2(\tau) = p_2^{SLII} - b_1 y_1^{SLII}\left[\frac{x_1^f}{y_1^f}\right]\left\{\frac{a_{11}}{(b_1 x_2^f)^2} + \left[q_1^{SLII}\left[\frac{y_1^f}{x_1^f}\right] + \frac{a_{11}}{b_2 x_2}\right](\tau - \tau^{II}_{SL}) - \frac{a_{11}}{(b_1 x_2^f)^2} e^{b_1 x_2^f (\tau - \tau^{II}_{SL})}\right\},$$

$$q_1(\tau) = q_1^{SLII} e^{-b_1 x_2^f (\tau - \tau^{II}_{SL})} - \left[\frac{x_1^f}{y_1^f}\right]\frac{a_{11}}{b_1 x_2^f}\left\{1 - e^{-b_1 x_2^f (\tau - \tau^{II}_{SL})}\right\},$$

$$q_2(\tau) = q_2^{SLII} - a_{12}\left[\frac{x_1^f}{y_1^f}\right](\tau - \tau^{II}_{SL}).$$

E6. $$P^{I}_{A2}: \begin{cases} u^*(\tau) = 1 \\ \qquad\qquad \text{for } \tau_v \leqslant \tau \leqslant \tau^{I}_{u'}. \\ v^*(\tau) = 0 \end{cases}$$

Results are similar to those for P^I_{A1} above in E4.

E7. $$P^I_{A3}: \begin{cases} u^*(\tau) = 1 \\ \qquad\qquad \text{for } \tau_v^* \leqslant \tau \leqslant \tau_u^I. \\ v^*(\tau) = 0 \end{cases}$$

Results are similar to those of P^I_{A1} above in E4.

E8. $$P^I_{S2}: \begin{cases} u^*(\tau) = b_2/(b_1+b_2)\cdot(1 - q_2 y_2/(p_2 x_2)) = u_S^* \\ \qquad\qquad \text{for } \tau_v^* \leqslant \tau \leqslant \tau_u^{-I}. \\ v^*(\tau) = 0 \end{cases}$$

As usual, we have that $S_v(\tau) > 0$ for $\tau > \tau_v$. In order for a U-singular subarc to be possible for $\tau \leqslant \tau_v^*$ the following condition must hold at $\tau = \tau_v^{*-}$:

$$b_1 p_2(\tau_v^{*-}) x_2(\tau_v^*) \geqslant b_2(-q_2(\tau_v^{*-})) y_2(\tau_v^*).$$

Also, on P^I_{S2} we have†

$$\frac{dx_1}{d\tau} = a_{11} x_1 y_1 \text{ with } x_1(\tau=\tau_v^*) = x_1^{v*},$$

$$\frac{dx_2}{d\tau} = a_2 y_2 \text{ with } x_2(\tau=\tau_v^*) = x_2^{v*},$$

$$\frac{dy_1}{d\tau} = u_S^* b_1 x_2 y_1 \text{ with } y_1(\tau=\tau_v^*) = y_1^{v*},$$

$$\frac{dy_2}{d\tau} = (1-u_S^*) b_2 x_2 y_2 \text{ with } y_2(\tau=\tau_v^*) = y_2^{v*},$$

and

$$\frac{dp_1}{d\tau} = -a_{11} y_1 p_1 \text{ with } p_1(\tau=\tau_v^*) = p_1^{v*},$$

$$\frac{dp_2}{d\tau} = -b_1 y_1 q_1 \text{ with } p_2(\tau=\tau_v^*) = p_2^{v*},$$

$$\frac{dq_1}{d\tau} = -a_{11} \frac{x_1^f}{y_1^f} - u_S^* b_1 x_2 q_1 \text{ with } q_1(\tau=\tau_v^*) = q_1^{v*},$$

$$\frac{dq_1}{d\tau} = -a_1 p_2 - (1 - u_S^*) b_2 x_2 q_2 \text{ with } q_2(\tau=\tau_v^*) = q_2^{v*}.$$

E9. $$P^{II}_{B1}: \begin{cases} u^*(\tau) = 1 \\ \qquad\qquad \text{for } 0 \leqslant \tau \leqslant \tau_u \text{ with } \dfrac{y_1^f}{y_2^f} < \left[\dfrac{y_1^f}{y_2^f}\right]^*. \\ v^*(\tau) = 1 \end{cases}$$

†A further discussion of the continuity of the adjoint variables is to be found in Section 5.2.4 of the main text.

τ_u is the smallest positive root of

$$\left[b_1 - \frac{a_{11}y_1^f}{x_2^f}\right] - a_{12}b_2y_2^f\tau_u + \frac{a_{11}y_1^f}{x_2^f}\, e^{b_1x_2^f\tau_u} = 0.$$

It should be noted that $\frac{\partial\tau_u}{\partial r} > 0$, where $r = y_1^f/y_2^f$. It may be shown that for $0 < \frac{y_1^f}{y_2^f} < \left(\frac{y_1^f}{y_2^f}\right)^*$,

$$\frac{\left(\frac{b_1}{b_2}\right)}{a_{12}y_2^f} < \tau_u < \tau_u^*,$$

where the determination of τ_u^* is given in E1. We also have that $\tau_u(r_1) < \tau_u(r_2)$ for $r_1 < r_2$ (x_1^f and x_2^f held constant). The solutions to the state and adjoint equations are the same as those for P_{S0}^{II} given previously. Let $S_v(\tau{=}\tau_u) = S_v^u$, $p_1(\tau{=}\tau_u) = p_1^u$, etc.

E10.
$$P_{B2}^{II}: \begin{cases} u^*(\tau) = 0 \\ \qquad\qquad \text{for } \tau_u \leqslant \tau \leqslant \tau_v. \\ v^*(\tau) = 1 \end{cases}$$

It follows that for all $\tau < \tau_u$ we have $S_u(\tau) < 0$ and $\frac{y_1}{y_2}(\tau) < a_{12}b_2/(a_{11}b_1)$. τ_v is the smallest positive root of $S_v(\tau{=}\tau_v) = 0$, where $S_v(\tau)$ is given by

$$S_v(\tau) = S_v^u + a_2b_2y_2^u\left(\frac{x_1^f}{y_1^f}\right)\left\{\frac{a_{12}}{(b_2x_2^f)^2} + \left[q_2^u\left(\frac{y_1^f}{x_1^f}\right) + \frac{a_{12}}{b_2x_2^f}\right](\tau-\tau_u) - \frac{a_{12}}{(b_2x_2^f)^2}\exp\left[b_2x_2^f(\tau-\tau_u)\right]\right\}.$$

Also, on P_{B2}^{II} we have

$$x_1(\tau) = x_1^u \exp\left\{a_{11}y_1^u(\tau-\tau_u) + \frac{a_{12}y_2^f}{b_2x_2^f}\left[e^{b_2x_2^f(\tau-\tau_u)} - 1\right]\right\},$$

$$x_2(\tau) = x_2^f,$$

$$y_1(\tau) = y_1^u,$$

$$y_2(\tau) = y_2^f \exp\left\{b_2x_2^f(\tau-\tau_u)\right\},$$

and

$$p_1(\tau) = p_1^u \exp\left\{-a_{11}y_1^u(\tau-\tau_u) - \frac{a_{12}y_2^f}{b_2x_2^f}\left[e^{b_2x_2^f(\tau-\tau_u} - 1\right]\right\} \text{ with } p_1^u = \frac{1}{x_1^u}\left(\frac{x_1^f}{y_1^f}\right),$$

$$p_2(\tau) = p_2^u - b_2y_2^f\left(\frac{x_1^f}{y_1^f}\right)\left\{\frac{a_{12}}{(b_2x_2^f)^2} + \left[q_2^u\left(\frac{y_1^f}{x_1^f}\right) + \frac{a_{12}}{b_2x_2^f}\right](\tau-\tau_u) - \frac{a_{12}}{(b_2x_2^f)^2}\, e^{b_2x_2^f(\tau-\tau_u)}\right\},$$

$$q_1(\tau) = q_1^u - a_{11}\left[\frac{x_1^f}{y_1^f}\right](\tau-\tau_u),$$

$$q_2(\tau) = \left[\frac{x_1^f}{y_1^f}\right]\left\{-\frac{a_{12}}{b_2x_2^f} + \left[\frac{a_{12}}{b_2x_2^f} + q_2^u\left[\frac{y_1^f}{x_1^f}\right]\right] e^{-b_2x_2^f(\tau-\tau_u)}\right\}.$$

E11. $$P_{B3}^{II}: \begin{cases} u^*(\tau) = 0 \\ \qquad \text{for } \tau_{SL}^{II} \leqslant \tau \leqslant \tau_v. \\ v^*(\tau) = 1 \end{cases}$$

τ_v is the smallest positive root of $S_v(\tau=\tau_v) = 0$, where

$$S_v(\tau) = S_v^{SLII} + a_2b_2y_2^{SLII}\left[\frac{x_1^f}{y_1^f}\right]\left\{\frac{a_{12}}{(b_2x_2^f)^2} + \left[q_2^{SLII}\left[\frac{y_1^f}{x_1^f}\right] + \frac{a_{12}}{b_2x_2^f}\right](\tau-\tau_{SL}^{II}) - \frac{a_{12}}{(b_2x_2^f)^2}\, e^{b_2x_2^f(\tau-\tau_{SL}^{II})}\right\}.$$

Again, an upper bound on τ_v is given by $a_{12}/(a_2b_1)$. It may be shown that $S_u(\tau) < 0$ for all $\tau > \tau_{SL}^{II}$. Also, on P_{B3}^{II} we have

$$x_1(\tau) = x_1^{SLII}\exp\left\{a_{11}y_1^{SLII}(\tau-\tau_{SL}^{II}) + \frac{a_{12}y_2^{SLII}}{b_2x_2^f}\left[e^{b_2x_2^f(\tau-\tau_{SL}^{II})} - 1\right]\right\},$$

$$x_2(\tau) = x_2^f,$$

$$y_1(\tau) = y_1^{SLII},$$

$$y_2(\tau) = y_2^{SLII}\exp\left\{b_2x_2^f(\tau-\tau_{SL}^{II})\right\},$$

and

$$p_1(\tau) = p_1^{SLII}\exp\left\{-a_{11}y_1^{SLII}(\tau-\tau_{SL}^{II}) - \frac{a_{12}y_2^{SLII}}{b_2x_2^f}\left[e^{b_2x_2^f(\tau-\tau_{SL}^{II})} - 1\right]\right\} \text{ with } p_1^{SLII} = \frac{1}{x_1^{SLII}}\left[\frac{x_1^f}{y_1^f}\right],$$

$$p_2(\tau) = p_2^{SLII} - b_2y_2^{SLII}\left[\frac{x_1^f}{y_1^f}\right]\left\{\frac{a_{12}}{(b_2x_2^f)^2} + \left[q_2^{SLII}\left[\frac{y_1^f}{x_1^f}\right] + \frac{a_{12}}{b_2x_2^f}\right](\tau-\tau_{SL}^{II}) - \frac{a_{12}}{(b_2x_2^f)^2}\, e^{b_2x_2^f(\tau-\tau_{SL}^{II})}\right\},$$

$$q_1(\tau) = q_1^{SLII} - a_{11}\left[\frac{x_1^f}{y_1^f}\right](\tau-\tau_{SL}^{II}),$$

$$q_2(\tau) = \left[\frac{x_1^f}{y_1^f}\right]\left\{-\frac{a_{12}}{b_2x_2^f} + \left[\frac{a_{12}}{b_2x_2^f} + q_2^{SLII}\left[\frac{y_1^f}{x_1^f}\right]\right] e^{-b_2x_2^f(\tau-\tau_{SL}^{II})}\right\}.$$

(E12). $$P_{B4}^{I}: \begin{cases} u^*(\tau) = 0 \\ v^*(\tau) = 0 \end{cases} \quad \text{for } \tau_v^* \leqslant \tau.$$

It may be shown that $S_u(\tau) < 0$ and $S_v(\tau) < 0$ for all $\tau > \tau_v^*$. Also on P_{B4}^I we have

$$x_1(\tau) = x_1^{v^*} \exp\{a_{11} y_1^{v^*} (\tau - \tau_v^*)\},$$

$$x_2(\tau) = \begin{cases} \sqrt{(x_2^{v^*})^2 - \frac{2a_2}{b_2} y_2^{v^*}} \coth(-A(\tau - \tau_v^*) + B) \text{ for } \frac{b_2}{2} (x_2^{v^*})^2 > a_2 y_2^{v^*}, \\ x_2^{v^*} / \left\{1 - \frac{b_2}{2} x_2^{v^*} (\tau - \tau_v^*)\right\} \text{ for } \frac{b_2}{2} (x_2^{v^*})^2 = a_2 y_2^{v^*}, \\ \sqrt{\frac{2a_2}{b_2} y_2^{v^*} - (x_2^{v^*})^2} \tan(C(\tau - \tau_v^*) + D) \text{ for } \frac{b_2}{2} (x_2^{v^*})^2 < a_2 y_2^{v^*}, \end{cases}$$

$$y_1(\tau) = y_1^{v^*},$$

$$y_2(\tau) = \begin{cases} \left(\frac{b_2 (x_2^{v^*})^2}{2a_2} - y_2^{v^*}\right) / \sinh^2(-A(\tau - \tau_v^*) + B) \text{ for } \frac{b_2}{2} (x_2^{v^*})^2 > a_2 y_2^{v^*}, \\ y_2^{v^*} / \left\{1 - \frac{b_2}{2} x_2^{v^*} (\tau - \tau_v)\right\}^2 \text{ for } \frac{b_2}{2} (x_2^{v^*})^2 = a_2 y_2^{v^*}, \\ \left(y_2^{v^*} - \frac{b_2}{2a_2} (x_2^{v^*})\right) / \cos^2(C(\tau - \tau_v^*) + D) \text{ for } \frac{b_2}{2} (x_2^{v^*})^2 < a_2 y_2^{v^*}, \end{cases}$$

where

$$A = \frac{b_2}{2} \sqrt{(x_2^{v^*})^2 - 2\frac{a_2}{b_2} y_2^{v^*}},$$

$$B = \coth^{-1}\left[\frac{x_2^{v^*}}{\sqrt{(x_2^{v^*})^2 - \frac{2a_2}{b_2} y_2^{v^*}}}\right],$$

$$C = \frac{b_2}{2} \sqrt{\frac{2a_2}{b_2} y_2^{v^*} - (x_2^{v^*})^2},$$

$$D = \tan^{-1}\left[\frac{x_2^{v^*}}{\sqrt{\frac{2a_2}{b_2} y_2^{v^*} - (x_2^{v^*})^2}}\right]$$

and

$$p_1(\tau) = p_1^{v^*} \exp\left\{-a_{11} y_1^{v^*}(\tau - \tau_v^*)\right\} \text{ with } p_1^{v^*} = \frac{1}{x_1^{v^*}}\left[\frac{x_1^f}{y_1^f}\right],$$

$$\frac{dp_2}{d\tau} = -b_2 y_2 q_2 \text{ with } p_2(\tau = \tau_v^*) = p_2^{v^*},$$

$$q_1(\tau) = q_1^{v^*} - a_{11}\left[\frac{x_1^f}{y_1^f}\right](\tau - \tau_v^*),$$

$$\frac{dq_2}{d\tau} = -a_2 p_2 - b_2 x_2 q_2 \text{ with } q_2(\tau = \tau_v^*) = q_2^{v^*}.$$

We have not been able to develop solutions in terms of "elementary" functions to the equations for p_2 and q_2.

E13. $$P_{B2}^I \text{ and } P_{B3}^I: \begin{cases} v^*(\tau) = 0 \\ \qquad\qquad \text{for } \tau_v \leqslant \tau. \\ v^*(\tau) = 0 \end{cases}$$

Results are similar to those for $P_{B4}sul$ previously in E12.

E14. $$P_{B5}^I: \begin{cases} u^*(\tau) = 0 \\ \qquad\qquad \text{for } \tau_{SL}^I \leqslant \tau. \\ v^*(\tau) = 0 \end{cases}$$

Results are similar to those for P_{B4}^I previously in E12.

E15. $$P_{A4}^I: \begin{cases} v^*(\tau) = 1 \\ \qquad\qquad \text{for } \tau_{SL}^I \leqslant \tau \leqslant \tau_u^I. \\ v^*(\tau) = 0 \end{cases}$$

Results are similar to those for P_{A1}^I previously in E4.

REFERENCES

[1] Antosiewicz, H., "Analytic Study of War Games," Naval Research Logistics Quarterly *2*, 181-208 (1955).

[2] Bellman, R., and S. Dreyfus, "On a Tactical Air-Warfare Model of Mengel," Operations Research *6*, 65-78 (1958).

[3] Berkovitz, L., "A Variational Approach to Differential Games," in *Advances in Game Theory*, pp. 127-174. M. Dresher, L. Shapley, and A. Tucker (Eds.), (Princeton University Press, Princeton, New Jersey, 1964).

[4] Berkovitz, L., "Lectures on Differential Games," in *Differential Games and Related Topics*, pp. 3-45. H. Kuhn and G. P. Szego (Eds.), (North-Holland Publishing Co., Amsterdam, 1971).

[5] Berkovitz, L. and M. Dresher, "A Game Theory Analysis of Tactical Air War," Operations Research *7*, 599-620 (1959).

[6] Berkovitz, L., and M. Dresher, "Allocation of Two Types of Aircraft in Tactical Air War: A Game Theoretic Analysis," Operations Research *8*, 694-706 (1960).

[7] Berkovitz, L., and M. Dresher, "A Multimove Infinite Game with Linear Payoff," Pacific Journal of Mathematics *10*, 743-765 (1960).

[8] Bonder, S., "The Lanchester Attrition-Rate Coefficient," Operations Research *15*, 221-232 (1967).

[9] Bonder, S., "The Mean Lanchester Attrition Rate," Operations Research *18*, 179-181 (1970).

[10] Bonder, S., and R. Farrell (Eds.), "Development of Models for Defense Systems Planning," Report No. SRL 2147 TR 70-2 (U), Systems Research laboratory, The University of Michigan, Ann Arbor, Michigan (Sept. 1970).

[11] Brackney, H., "The Dynamics of Military Combat," Operations Research *7*, 30-44 (1959).

[12] Bryson, A., and Y. C. Ho, *Applied Optimal Control*, (Blaisdell Publishing Co., Waltham, Massachusetts, 1969).

[13] Dresher, M., *Games of Strategy: Theory and Applications*, (Prentice-Hall, Englewood Cliffs, New Jersey, 1961).

[14] Freidman, A., *Differential Games*, (Wiley-Interscience, New York, 1971).

[15] Fulkerson, D., and S. Johnson, "A Tactical Air Game," Operations Research *5*, 704-712 (1957).

[16] Giamboni, L., A. Mengel, and R. Dishington, "Simplified Model of a Symmetric Tactical Air War," The RAND Corporation, RM-711 (August 1951).

[17] Harris, K., and L. Wegner, "Tactical Airpower in NATO Contingencies: A Joint Air-Battle/Ground-Battle Model (TALLY/TOTEM)," The RAND Corporation, R-1194-PR, (May 1974).

[18] Ho, Y. C., "Review of the book *Differential Games* by R. Isaacs," IEEE Transaction on Automatic Control Vol. AC-10, 501-503 (1965).

[19] Ho, Y. C., "Differential Games, Dynamic Optimization, and Generalized Control Theory," Journal of Optimization Theory and Applications *6*, 179-209 (1970).

[20] Isaacs, R., "Differential Games IV: Mainly Examples," The RAND Corporation, RM-1486 (March 1955).

[21] Isaacs, R., *Differential Games*, (John Wiley & Sons, Inc., New York, 1965).

[22] Kawara, Y., " An Allocation Problem of Fire Support in Combat as a Differential Game," Operations Research *21*, 942-951 (1973).

[23] Kelley, H., "Second Variation Test for Singular Extremals," AIAA Journal *2*, 1380-1382 (1964).

[24] Kelley, H., R. Kopp, and H. Moyer, "Singular Extremals," in *Topics in Optimization*, G. Leitman (Ed.) (Academic Press, New York, 1967).

[25] Lansdowne, Z., G. Dantzig, R. Harvey, and R. McKnight, "Development of an Algorithm to Solve Multi-State Games," Control Analysis Corporation, Palo Alto, California (May 1973).

[26] McNicholas, R., and F. Crane, "Guide to Fire Support Mix Evaluation Techniques, Volume I: The Guide and Appendices A and B," Stanford Research Institute, Menlo Park, California (March 1973).

[27] Morse, P., and G. Kimball, *Methods of Operations Research*, (M.I.T. Press, Cambridge, Massachusetts, 1951).

[28] Pugh, G., "Theory of Measures of Effectiveness for General-Purpose Military Forces: Part II. Lagrange Dynamic Programming in Time-Sequential Combat Games," Operations Research *21*, 886-906 (1973).

[29] Pugh, G., and J. Mayberry, "Analysis of General-Purpose Force Procurement," Lambda Corporation, Paper 59 (August 1971).

[30] Pugh, G., and J. Mayberry, "Theory of Measures of Effectiveness for General-Purpose Military Forces: Part I. A Zero-Sum Payoff Appropriate for Evaluating Combat Strategies," Operations Research *21*, 867-885 (1973).

[31] Taylor, J., "Target Selection in Lanchester Combat: Linear-Law Attrition Process," Naval Research Logistics Quarterly *20*, 673-697 (1973).

[32] Taylor, J., "Lanchester-Type Models of Warfare and Optimal Control," Naval Research Logistics Quarterly *21*, 79-106 (1974).

[33] Taylor, J., "Some Differential Games of Tactical Interest," Operations Research *22*, 304-317 (1974).

[34] Taylor, J., and G. Brown, "Canonical Methods in the Solution of Variable-Coefficient Lanchester-Type Equations of Modern Warfare," Operations Research *21*, 44-69 (1976).

[35] Taylor, J., and S. Parry, "Force-Ratio Considerations for Some Lanchester-Type Models of Warfare," Operations Research *23*, 522-533 (1975).

[36] USAF Assistant Chief of Staff, Studies and Analysis, "Methodology for Use in Measuring the Effectiveness of General Purpose Forces, SABER GRAND (ALPHA)," (March 1971).

[37] Weiss, H., "Lanchester-Type Models of Warfare," in *Proceedings of the First International Conference Operational Research*, pp. 82-98, (Operations Research Society of America, Baltimore 1957).

[38] Weiss, H., "Some Differential Games of Tactical Interest and the Value of a Supporting Weapon System," Operations Research *7*, 180-196 (1959).

ON THE MANIPULATION OF TRANSFER PRICES IN A STATIC ENVIRONMENT

John P. Bonin

Wesleyan University
Middletown, Connecticut

ABSTRACT

In a static environment, J. Hirschleifer's marginal cost solution to the transfer pricing problem is commonly accepted as analytically correct. However, actual pricing practice within Western corporations and socialist-planned economies generally deviates from marginal cost pricing. Some form of average cost pricing is more commonly chosen. Recently in this journal, H. Enzer has claimed to show that some form of average cost pricing is indeed the analytically correct solution to the transfer pricing problem when choice of technique and manipulation are allowed. Enzer claims that optimal decisions made by each of two divisions according to their individual self-interests are made compatible with overall firm optimization when the transfer price assigned to the internally-transferred commodity is any form of average cost.

We show that the marginal cost solution is correct for Enzer's problem in the absence of manipulation by either division. Indeed, this was all that Hirschleifer claimed. In the process, we uncover a fundamental mathematical error in Enzer's argument. When manipulation of the transfer price by divisions is allowed, we demonstrate the faults with Enzer's average cost solution and conclude Hirschleifer's original statements on manipulation to be correct even in Enzer's environment. A final section briefly indicates the importance to the transfer pricing problem of a growing body of economic literature on incentive structures.

1. INTRODUCTION: THE STRUCTURE OF THE PROBLEM

In a recent issue of this journal [4], H. Enzer claims to demonstrate mathematically that J. Hirschleifer's [9] marginal cost solution to the transfer pricing problem is inappropriate. In his paper, Enzer argues that some form of average cost is the theoretically correct transfer price. Since the textbooks on managerial economics that we consulted accept the analytical validity of Hirschleifer's work,* Enzer's result, if correct, would be a significant criticism of accepted theory. Interestingly enough, Enzer would have also provided analytical justification for institutionally accepted pricing practice in both Western corporations and socialist planned economies.† The purpose of this note is to show that Enzer's result is theoretically invalid. In

*Cf. Haynes and Henry [8], pp. 413-424, and McGuigan and Moyer [14], pp. 349-360. Gupta and Cozzolino [6], p.32, provide an interesting numerical example of marginal cost transfer pricing in a retractable pen company.

†Bornstein [3] is a discussion of socialist pricing practice while Kyn, Sekerka, and Hejl [11] analyze socialist pricing theory. McGuigan and Moyer [14], pp. 354-360, consider corporate pricing practices.

the process, we demonstrate the fundamental mathematical error in Enzer's paper. A concluding section discusses the relevance of work on incentive structures to the transfer pricing problem.

The basic nature of any decentralization problem is to structure incentives so that myopic optimization by individual units generates optimization of aggregate goals. The particular problem of transfer pricing involves a firm composed of two divisions, or profit centers. The firm and its divisions have complete deterministic knowledge of all cost, demand, and production relationships. One assumes that each division maximizes its own profit while the firm attempts to maximize total profit. The divisions are vertically integrated, since the output of one is used as an input by the other. In the simplest case, the transfer of this commodity takes place internally in the absence of external markets. Final output and any other inputs are transacted on competitive markets so that these prices are parametric to the firm (and, hence, the division). To complete the specification of the environment for the transfer pricing problem, it is necessary to consider the decision-making structure.

Fundamental to the determination of a transfer pricing solution to the decentralization problem is an assignment of decision variables to various agents. In Enzer's paper, the production division chooses the level of usage for two inputs (χ_1, χ_2), and the distribution division chooses input levels of the intermediate commodity (u), produced by the other division, along with another input (χ_3). The decision environment in the initial Hirschleifer article is simpler. Here the distribution division of a firm incurs a cost which varies with the level of the intermediate commodity marketed. The production department incurs production costs which vary with output. In both environments, the transfer price is to be determined so that the output level of the intermediate commodity which maximizes total profit will also be profit maximizing for each division.

Hirschleifer demonstrates that the firm should set a transfer price equal to the marginal cost of producing the output which maximizes total profits. If each division maximizes its own profit using this transfer price to evaluate interdivisional transactions, the optimal output level for each division will be equal to the one which maximizes total profit. Consequently, marginal cost pricing solves the decentralization problem. In the following section, Enzer's complicating addition of the choice of technique by each division is shown to be a red herring. The Hirschleifer result reported above is derived in such an environment.

Crucial to the solution of this decentralization problem is the stipulation that the firm, an agent separate from each division, imposes a transfer price which is taken to be a parameter in the individual division's maximization problem. Given that the firm has full information concerning costs and demands, this position seems tenable. However, Enzer seemingly allows the possibility that each division might consider the influence of its decisions on the transfer price and attempt to manipulate the price in its own self-interest.* In his paper, Hirschleifer considers the case where the distribution division anticipates a transfer price set equal to marginal production cost at the level of output it chooses. This division would then find as optimal a level of output smaller than that which maximizes total profit. At the same time, Hirschleifer indicates that, if both divisions anticipate the transfer pricing rule, a bilateral monopoly develops.† Consequently, dual manipulation yields a bargaining problem with no consistent price-decentralizable solution.

*Refer to Enzer [4], p. 379, first-order conditions number one for division one and number two for division two. Notice that r_u is not considered parametric in either.

†See Hirschleifer [9], p. 310, for a discussion of manipulation.

Enzer uses this manipulation argument to suggest that the firm must fix the level of operations in Hirschleifer's problem to solve the decentralization problem. In what follows, the manipulation problem is shown to be present, both in Enzer's formulation and in his incorrect average cost solution. When dual manipulation of the sort discussed above is assumed, a profit-maximizing production division produces more of the intermediate commodity than a profit-maximizing distribution division is willing to accept. We also demonstrate that the introduction of an external market for this intermediate commodity is not the solution to bilateral bargaining. Indeed, by pointing out Enzer's fundamental mathematical error, we cast severe doubt on the validity of his entire discussion, since all the cases he considers depend on this incorrect assertion.

In the literature, transfer prices are allotted two functions, that of resource allocation and of performance evaluation. In the spirit of Hirschleifer's article and to correct Enzer's mistaken assertions, we concentrate in the next section on the resource allocation problem (i.e., marginal conditions). Enzer's analysis addresses this role but his exposition seems to confuse the two functions. Indeed, when participating agents have perfect knowledge of all demand, cost and production conditions, the absolute level of either division's profits calculated by using the transfer price provides no new information for performance evaluation. However, once Enzer's assumption of perfect knowledge is dropped, the transfer price problem comes under the broad rubric of resource allocation with incomplete information. A concluding section briefly discusses some of the recent interesting contributions to this topic in the economic literature. Several results on designing incentive structures to combat manipulation and to generate "truthful" responses under incomplete knowledge are discussed. Further work on the theory of transfer pricing should at least take cognizance of this new research in economics.

2. THE ANALYTICAL PROOF

Initially, we demonstrate the validity of Hirschleifer's transfer price solution when a choice of technique by both divisions is assumed but manipulation of the transfer price by either division is not allowed. The firm then stipulates that any transactions of the intermediate commodity between divisions will take place at a transfer price equal to the marginal cost of producing the output level of this commodity which maximizes total firm profits. Note that this price level is set independent of any maximization decisions made by the individual divisions. On the other hand, when either division can manipulate the transfer price by altering its own decisions, no consistent pricing solution which harmonizes the conflicting interests of the two divisions can be derived. Enzer's average cost pricing schemes are subject to the same criticism he levels at Hirschleifer's marginal cost scheme. Enzer's mathematical error is shown to involve equating several partial derivatives which should be evaluated at different realizations of their arguments.

Following Enzer's notation, let

$u = f(\chi_1, \chi_2)$ represent production of the intermediate commodity,
$q = g(\chi_3, u)$ represent production of the final product, and
$\bar{p}, \bar{r}_1, \bar{r}_2, \bar{r}_3$, be the respective price parameters for final output and the three inputs bought in the market.

The firm wishes to solve

Max: $\pi \equiv \bar{p}g(\chi_3, u) - \bar{r}_1\chi_1 - \bar{r}_2\chi_2 - r_3\chi_3$,
where $u = f(\chi_1, \chi_2)$.

Rather than use Lagrangian analysis, we substitute for u and obtain the following three first-order conditions:*

*We define $f_i = (\partial f(\chi_1, \chi_2)/\partial\chi_i)$ for $i = 1, 2$, $g_u = (\partial g(\chi_3, u)/\partial u)$, and $g_3 = (\partial g(\chi_3, u)/\partial\chi_3)$.

(1) $$\partial\pi/\partial\chi_1 = \bar{p}g_u f_1 - \bar{r}_1 = 0$$

(2) $$\partial\pi/\partial\chi_2 = \bar{p}g_u f_2 - \bar{r}_2 = 0.$$

(3) $$\partial\pi/\partial\chi_3 = \bar{p}g_3 - \bar{r}_3 = 0.$$

We can solve (3) for χ_3 as a function of u, call it $\chi_3^*(u)$ and substitute in (1) and (2). Necessary conditions for a maximum for the firm's profits are then*

(1)′ $$\partial\pi/\partial\chi_1 = \bar{p}\left(g_3 \frac{d\chi_3^*}{du} + g_u\right) f_1 - \bar{r}_1 = 0$$

(2)′ $$\partial\pi/\partial\chi_2 = \bar{p}\left(g_3 \frac{d\chi_3^*}{du} + g_u\right) f_2 - \bar{r}_2 = 0.$$

The production division chooses χ_1 and χ_2 to maximize its own profit, using the transfer price r_u to evaluate output; i.e.,

Max: $\pi_1 \equiv r_u f(\chi_1, \chi_2) - \bar{r}_1\chi_1 - \bar{r}_2\chi_2$,
with regard to χ_1, χ_2

First-order conditions are

(4) $$\frac{\partial\pi_1}{\partial\chi_1} = \left(r_u + f(\chi_1, \chi_2)\frac{\partial r_u}{\partial u}\right) f_1 - \bar{r}_1 = 0$$

(5) $$\frac{\partial\pi_1}{\partial\chi_2} = \left(r_u + f(\chi_1, \chi_2)\frac{\partial r_u}{\partial u}\right) f_2 - \bar{r}_2 = 0$$

Let us denote as u_P the optimal output level which results from χ_1^*, χ_2^* solving (4) and (5), i.e., $u_P = f(\chi_1^*, \chi_2^*)$.

The distribution division chooses u, here designated as u_D, and χ_3 to maximize profits by using the transfer price to evaluate the cost of the intermediate commodity used. The problem is

Max: $\pi_2 \equiv \bar{P}g(\chi_3, u) - \bar{r}_3\chi_3 - r_u u$,
with regard to u, χ_3

First-order conditions are

(6) $$\frac{\partial\pi_2}{\partial\chi_3} = \bar{p}g_3 - \bar{r}_3 = 0$$

(7) $$\frac{\partial\pi_2}{\partial u} = \bar{p}g_u - r_u - u\frac{\partial r_u}{\partial u} = 0.$$

Using (6) to solve $\chi_3^*(u)$ and substituting in (7), we obtain a single necessary condition

(7)′ $$\bar{p}\left(g_3 \frac{d\chi_3^*(u)}{du} + g_u\right) - r_u - u\,\partial r_u/\partial u = 0.$$

The decentralization problem is solved when u_p from (4) and (5), u_D from (7)′, and u^* from (1)′ and (2)′ are all equal. Algebraic operations on Enzer's conditions generate (1), (2),

*Following Enzer, we assume sufficient concavity of production functions to ensure that first-order conditions represent a global maximum.

and (3) as the first-order conditions for the firm's maximization problem, and (4) and (5) as those for a profit-maximizing division 1. Also, (6) and (7) are identical to the necessary conditions in Enzer for division 2 to profit-maximize. Consequently, our problem is formally equivalent to Enzer's problem. Our simplification helps to clarify the important aspects of the analysis.

Firstly, assume that the firm can impose a transfer price so that transactions between divisions will take place only at that price. This implies $\frac{\partial r_u}{\partial u} = 0$ in (4), (5), and (7)′. How should r_u^* be determined? It is clear from a comparison of (4) and (5) with (1)′ and (2)′ that the choice of technique which maximizes both the firm's profits and division 1 profits will be attained for any specification of r_u, the transfer price. Hence, the choice-of-technique complication is simply a red herring. However, the optimal amount of the intermediate commodity produced and used from the viewpoint of maximizing total firm profits (denote this output u^*) is generated if and only if r_u^* is set by the firm to solve

$$r_u^* = \bar{p}(g_3 d\chi_3(u^*)/du^* + g_u^*). \tag{8}$$

Then, maximizing π and maximizing π_1 will yield the same level of operations, i.e., $u^* = u_P$. Substitution of (8) into (7)′ yields $u_D = u^*$ when π_2 is maximized. Consequently, the decentralization problem is solved; i.e., $u^* = u_P = u_D$, when the firm imposes the transfer price which solves (8) upon all transactions between divisions.

In economic terms, (8) is interpreted to mean that the transfer price should equal the value marginal product to the distribution division of using the amount of the intermediate commodity u^* which maximizes total firm profit. However, from (1)′ or (2)′, the right-hand side of (8) can also be shown to equal the marginal cost of producing this optimal (from the firm's viewpoint) level of the intermediate commodity u^*. Note that from the definition of the marginal cost of producing u, $\frac{dC}{du} = \frac{dC_1}{du} = \frac{\bar{r}_1}{f_1} = \frac{\bar{r}_2}{f_2}$.† Consequently, (8) generates the Hirschleifer result that in the absence of manipulation, the transfer price r_u^* should equal the marginal cost of producing u, evaluated at that output level which maximizes total firm profits, i.e., u^*. Introducing a choice of technique for each division, though cluttering the analysis, does not change the analytical result.

How then, did Enzer conclude that the correct transfer price is some variant of average cost? If the result is true, it must depend on his allowing manipulation of the transfer price by divisions. In the above analysis, suppose $\frac{\partial r_u}{\partial u} \neq 0$; i.e., the transfer price depends on the level of the transaction between the two divisions. Either a transfer-pricing rule defines this derivative or the divisions form expectations about how their actions affect r_u. In Hirschleifer's example referred to above $\frac{\partial r_u}{\partial u_D} > 0$ because the distribution division knows that the transfer price r_u would be set equal to the marginal cost of producing the level of the intermediate commodity it demanded.‡ However, the production division did not expect to influence the transfer price so that $\frac{\partial r_u}{\partial u_P} = 0$. Therefore, (1)′, (2)′, (4), and (5) are unchanged. However, (7)′ is now less than zero at u, given r_u^*. Consequently, $u_D < u^* = u_P$, as Hirshleifer suggests.

†The first equality represents the identity between the marginal cost of producing the commodity both for the firm and the production division. The latter equalities follow from partial differentiation of the definition of the cost of producing u, i.e., $C = C_1 = r_1\chi_1 + r_2\chi_2$, noting that $\partial\chi_i/\partial u = 1/f_i$.

‡We are assuming that marginal cost increases with the production of u in the relevant range.

Suppose both divisions expect the transfer price to respond positively to the level of the intermediate commodity transacted because of the marginal cost pricing rule. Then, (4) and (5) will be positive at u^*, given r_u^*. Consequently, $u_D < u^* < u_P$. If, for some reason, $\frac{\partial r_u}{\partial u} < 0$ were expected by both divisions, the inequalities would be reversed. In any case, allowing the transfer price to be influenced by the level of the intermediate commodity transferred destroys decentralization by the pricing mechanism. The conflicting self-interests of the two profit-maximizing divisions can not be harmonized. As Hirschleifer points out, the bilateral monopoly which results must be resolved through bargaining strategies.

The average cost solution, embraced by Enzer, implies $\frac{\partial r_u}{\partial u} \neq 0$ since the transfer price is linearly related to average cost. Hence, in the presence of manipulation, it falls victim to the same criticism. Indeed, if the transfer price were set equal to average production cost, the production division would earn zero profits on any level of output. Hence, all levels of production would be equally desirable for this division. The distribution division would want average production costs to be minimized, and hence its own profit maximized. Therefore, the cost-minimizing output level should be imposed on the production division. But this is precisely what Enzer objected to in the Hirschleifer problem.

Enzer's fundamental mathematical error involves equating several partial derivatives which are themselves functions of u, the level of the intermediate commodity transferred. The second equation (p. 380) reads

$$\frac{\partial(r_u u)}{\partial u} = -\lambda_1 = -\lambda = MC_1 = dC_1/du. \tag{9}$$

Each expression is a function of u, with λ_1 the marginal cost of producing the level of u which maximizes the production division's profits i.e., u_P, and λ the marginal production cost of u^*, the level which maximizes total profits. The left-hand side of (9) should be evaluated at u_D, the level which maximizes the distribution division's profits.* Therefore, (9) is true only if the partial derivatives involved are evaluated at the same argument, i.e., $u_D = u^* = u_P$. Consequently, Enzer assumes what he is trying to prove, i.e., that the decentralization problem is solvable by the pricing mechanism. Indeed, his specification is overdetermined, which accounts for the arbitrary constant of integration appearing in his price equation. Enzer does not solve the transfer pricing problem when manipulation is allowed.

We point out, in passing, that admitting an external market for the intermediate commodity also does not yield Enzer's result. Enzer's fundamental mathematical error is continued throughout this article; hence, his remaining discussions are all incorrect. As Hirschleifer suggests, when the external market is competitive, the transfer price must equal the market price. For any transfer price below the market price, the producing division will maximize profits by selling all of its output on the external market. At any transfer price above the market price, the distribution division will maximize profits by buying all of its input from the market. Hence, the only price at which transfers between divisions will take place is equal to the external competitive market price. The situation becomes more complex when external markets are not competitive. However, none of Hirschleifer's original insights nor the later corroborating literature are rendered incorrect by Enzer's paper. In the absence of manipulation, the correct

*The sentence below the equation states "...and Div. 2 will demand input u until the marginal revenue product Pg_u equals the marginal cost of producing u,..." This is true only if the LHS of (9), $\partial(r_u u)/\partial u$, is evaluated at u_D.

transfer price is marginal cost. On the other hand, manipulation evokes a monopolistic response from the manipulator, and destroys decentralization by the pricing mechanism.

3. DESIGNING INCENTIVE STRUCTURES TO ALLOCATE RESOURCES

Despite its technical faults, Enzer's paper has raised important issues in the practical application of transfer pricing to transactions within a capitalist corporation. Indeed, many such organizations and virtually all socialist planned economies use some variant of average cost to determine the price of intermediate commodities transferred internally. While attempting to simulate the desirable allocative consequences of competitive markets, these organizations have been forced to consider the other role of monetary variables. When production and market conditions are not known with perfect certainty, the level of profits is often a performance indicator.* Hence, when marginal cost pricing leads to negative profits (i.e., in the decreasing costs case), Western authors refer to a conflict between divisions over the correct mode of transfer pricing. Although O. Lange [12] argued that one advantage of socialism is that transactions can take place at marginal cost in this situation with the level of profits unimportant, his solution requires that some agent possess complete information about all important economic phenomena.† Consequently, when knowledge about the economic environment is incomplete, due to either dispersion within a hierarchical organization or stochastic elements, transfer prices are important as both resource allocators and performance evaluators.

An extensively studied area in the literature on decentralization under uncertainty is the team problem (cf. Marschak and Radner [13]). In a single-tier hierarchical organization, the center attempts to allocate a centrally held resource to individual independent divisions whose production relations are not vertically linked. Information is dispersed as each division is the sole agent possessing knowledge about its own production conditions, which may be stochastic. In a proper team, all agents wish to maximize total team profits. The object is to pass sufficient information to the center, given constraints on the capacity of information channels, to allow the solution of the complete optimization problem. An extremely interesting extension by Groves [5] constructs incentive schemes which force individual divisions, acting in their own self-interest and maximizing own profit only, to achieve the proper team solution in which total profit is maximized. Hence, Groves has designed an incentive structure within the team framework which generates complete optimization for the organization even when individual divisions pursue their own self-interest. Enzer's characterization of the transfer pricing problem indicates the importance of this problem in a vertically-integrated hierarchial organization.

In this spirit, when the firm does not have complete information about production conditions, the Hirschleifer solution developed in the previous section is unavailable. The firm can not calculate marginal or average cost relationships. Consequently, production information must be obtained from the divisions so that the firm can set the transfer price.‡ Why would either division respond truthfully rather than attempt to manipulate the transfer price in its own favor by responding in the monopolistic way described in the previous section?§ Such general problems in the incentive compatability of resource allocational mechanisms are presently being

*Cf. Hayek [7] for the classic statement of this point.

†Cf. Bergson [1] for a criticism of Lange's scheme on this point.

‡For Enzer's example, the firm needs information about production from both divisions in order to solve total profit maximization. In Hirschleifer's example, the firm needs production information only from the production division and cost information from the distribution division. To the economist, this is often equivalent information.

§Cf. Bergson [1] for a similar point about Lange.

researched by economists. L. Hurwicz [10] provides an excellent characterization of the problem and a review of the earlier literature. In Bonin [2], we have derived useful properties for a piecewise linear incentive structure in a single-tier hierarchical situation with dispersion of knowledge. The producing agent provides a truthful ex ante estimate of production conditions and subsequently strives for maximal ex post performance once the stochastic production variable is realized. The center is able to impose any degree of risk of plan underfulfillment on the producing agent by manipulating the parameters of the incentive scheme. We are presently working on embedding this incentive structure in a vertically-linked hierarchical environment. Results obtained should shed light on optimal incentive schemes for decentralized planning of intermediate transactions within both Western corporations and socialist economies.

ACKNOWLEDGMENTS

The author wishes to thank his colleague, T. Whitin, for calling this problem to his attention and for providing valuable comments on this paper.

BIBLIOGRAPHY

[1] Bergson, A., "Market Socialism Revisted," Journal of Political Economy, *75(5),* 655-673 (1967).

[2] Bonin, J. P., "On the Design of Managerial Incentive Structures in a Decentralized Planning Environment," American Economic Review, *66(4),* 682-687 (1976).

[3] Bornstein, M., "Soviet Price Theory and Policy," in M. Bornstein and D. Fusfeld (eds.), *The Soviet Economy: A Book of Readings, Fourth Edition,* (Richard Irwin, Inc. Homewood, Ill., 1974), pp. 85-117.

[4] Enzer, H., "The Static Theory of Transfer Pricing," Naval Research Logistics Quarterly, *22(2),* 375-389 (1975).

[5] Groves, T., "Incentives in Teams," Econometrica, *41(4),* 617-633 (1973).

[6] Gupta, S., and J. Cozzolino, Fundamentals of Operations Research for Management (Holden-Day San Francisco, 1974).

[7] Hayek, F., "The Use of Knowledge in Society," American Economic Review, *35(4),* 519-30 (1945).

[8] Haynes, W., and W. Henry, *Managerial Economics: Analysis and Cases, Third Edition,* (Business Publications, Inc., Dallas, 1974).

[9] Hirschleifer, J., "On the Economics of Transfer Pricing," Journal of Business, *29,* 172-184 (1956).

[10] Hurwicz, L., "The Design of Mechanisms for Resource Allocation," American Economic Review, *63(2),* 1-31 (May 1973).

[11] Kyn, O., B. Sekerka, and L. Hejl, "A Model for the Planning of Prices," in M. Bacharach and E. Malinvaud (eds.), *Activity Analysis in the Theory of Growth and Planning,* (St. Martin's Press, London, 1967) pp. 101-125.

[12] Lange, O., "On the Economic Theory of Socialism," in B. Lippincott (ed.), *On the Economic Theory of Socialism,* (The University of Minnesota Press, 1938) pp. 55-143.

[13] Marschak, J., and R. Radner, *Economic Theory of Teams,* (Yale University Press, New Haven, 1972).

[14] McGuigan, J., and R. Moyer, *Managerial Economics: Private and Public Sector Decision Analysis,* (Dryden Press, Hinsdale, Ill., 1975).

[15] Menge, S., "The Backward Art of Inter-divisional Transfer Pricing," Journal of Industrial Economics 9, 215-232 (1961).

A NOTE CONCERNING ASYMMETRIC GAMES ON GRAPHS

Alvin E. Roth

University of Illinois
Urbana, Illinois

ABSTRACT

We consider a class of asymmetric two-person games played on graphs, and characterize all the positions in the game.

In this note we consider a class of asymmetric two-person games in which the players alternately choose from a set of permissible moves, and the object of the game is to make the last move. We will characterize the positions in the game as winning, losing, or drawing from the point of view of each of the players.

Formally, let (V, A) be a directed graph, where V is the (finite) set of vertices, and $A \subset VxV$ is a set of directed arcs such that $A = A_1 \cup A_2$. The players take turns choosing vertices of the graph according to the rule* that if vertex v has just been chosen by player j, then player i $(i \neq j)$ may choose any vertex x such that $(v, x) \in A_i$. That is, player i may move along arcs in A_i. Player i loses the game (and his opponent wins) if it becomes his turn to move from a vertex v such that no arc (v, x) is in A_i.

If $A_1 = A_2$, the game is called "impartial" or "symmetric;" otherwise it is called "partial" or "asymmetric."† Most of the literature on games of this sort has concentrated on symmetric games,‡ but we will show that the vertices of an asymmetric game can be characterized in a natural way, which generalizes the results obtained for symmetric games [2].

In particular, we will characterize the set W_i of vertices that are *winning* for player i in the sense that if he chooses a vertex w in W_i then he can assure himself of eventually winning the game. Similarly, we will find the set L_i of vertices which are *losing* for player i, such that if he chooses a vertex v in L_i then he cannot prevent an eventual loss, and the set of *drawing* vertices D_i which are neither winning nor losing. Since the game is asymmetric, the resulting partition of the vertices is in general dependent on which of the two players is under consideration. Following Steinhaus [3] and Smith [4], we will also be interested in characterizing the (minimax) number of moves which remain from each vertex. (Steinhaus' interest in this question arose from the consideration of problems of naval pursuit.)

It will be convenient to consider, for every vertex x, the set of vertices from which player i can reach x, defined by $R_i(x) = \{v \in V | (v,x) \in A_i\}$. For every set of vertices S, denote the set

*The game starts when player 1 selects a vertex from some initial set I of vertices.

†An equivalent model (cf. Smith [4]) treats asymmetric games as having one set of arcs, but two sets of vertices.

‡With the notable exception of a recent book by J. H. Conway [1].

of vertices from which player i can reach some vertex in S by $R_i(S) = \bigcup_{x \in S} R_i(x)$, and let $U_i(S) = V - R_i(S)$ be the set of vertices from which the set S is unreachable by player i.

Note that $U_i(V)$ is the set of vertices from which player i can reach no other vertex: i.e. the set of vertices from which player i has no permissible moves. So player j ($j \neq i$) wins the game if he chooses a vertex v in $U_i(V)$.

For each positive integer n, define the sets of vertices B_n and C_n by $B_0 = C_0 = \phi$, $B_n = U_2(U_1(B_{n-1}))$, and $C_n = U_1(U_2(C_{n-1}))$. Observe that $C_1 = U_1(U_2(\phi)) = U_1(V)$, and $B_1 = U_2(V)$. So player 2 wins if he chooses a vertex in C_1, and player 1 wins by choosing a vertex in B_1.

The relationship between the sets B_n and C_n is given, for every n, by the following proposition:

PROPOSITION: (i) $B_{n-1} \subset B_n$, and $C_{n-1} \subset C_n$. (ii) $B_n \subset U_2(C_n)$, and $C_n \subset U_1(B_n)$.

PROOF: Observe that for $S, T, \subset V$, if $S \subset T$ then $U_i(S) \supset U_i(T)$, and $U_i(U_j(S)) \subset U_i(U_j(T))$, for $\{i, j\} = \{1, 2\}$. Also note that, when $n = 1$, both propositions (i) and (ii) hold (since for any $S \subset V$, $\phi \subset S$ and $U_i(V) \subset U_i(S)$).

Suppose that for some fixed n, it has been shown that $B_{n-1} \subset B_n$ and $C_{n-1} \subset C_n$. Then $B_n = U_2(U_1(B_{n-1})) \subset U_2(U_1(B_n)) = B_{n+1}$, and $C_n = U_1(U_2(C_{n-1})) \subset U_1(U_2(C_n)) = C_{n+1}$, and so proposition (i) is true for all n.

Suppose it has been shown for some n that $B_n \subset U_2(C_n)$ and $C_n \subset U_1(B_n)$. Then $U_1(B_n) \supset U_1(U_2(C_n)) = C_{n+1}$, and so $B_{n+1} = U_2(U_1(B_n)) \subset U_2(C_{n+1})$. The same argument shows that $C_{n+1} \subset U_1(B_{n+1})$, and so proposition (ii) is true for all n.

Note that proposition (i) together with the facts that V is finite implies that there is some integer k such that $B_k = B_{k+1} = U_2(U_1(B_k))$, and $C_k = C_{k+1} = U_1(U_2(C_k))$. Of course proposition (i) also implies that $B_k = \bigcup_n B_n \equiv B$ and $C_k = \bigcup_n C_n \equiv C$.

We can now characterize the sets of vertices which are winning, losing, and drawing for each player, as follows:

THEOREM: (i) $W_1 = B$, and $W_2 = C$. (ii) $L_1 = R_2(C)$, and $L_2 = R_1(B)$. (iii) $D_1 = U_2(C) - B$, and $D_2 = U_1(B) - C$.

PROOF: First we show that W_1 contains B; i.e. if player 1 picks a vertex in B, then he can make his subsequent choices so as to eventually win the game. Recall that $B = \bigcup_n B_n$, and suppose that player 1 picks a vertex b_n in B_n. If b_n is in B_1 then player 1 has won, otherwise player 2 picks a vertex x such that $b_n \in R_2(x)$. Since $B_n = U_2(U_1(B_{n-1}))$, it follows that $b_n \in R_2(U_1(B_{n-1}))$, and hence $x \in U_1(B_{n-1})$: i.e. $x \in R_1(B_{n-1})$. So no matter what vertex x is picked by player 2, player 1 will always be able to respond by choosing a vertex b_{n-1} in B_{n-1}. After at most n choices of this sort, player 1 picks a vertex b_1 in B_1, and wins. So W_1 contains B, and similar argument shows that W_2 contains C.

It is an immediate consequence that L_1 contains $R_2(C)$ and L_2 contains $R_1(B)$, since if player 2, for instance, picks a vertex in $R_1(B)$, then he cannot prevent player 1 from picking a

vertex in B and eventually winning. It only remains to show that $D_1 = U_2(C) - B$ and $D_2 = U_1(B) - C$: this will exhaust the set of vertices, and thus prove parts (i) and (ii), as well as (iii).

Suppose player 2 picks a vertex $v \in U_1(B) - C$. Then player 1 must choose a vertex w such that $v \in R_1(w)$, and so $w \in B$, since $v \in U_1(B)$. If player 1 chooses a vertex $w \in R_2(C)$, then we have seen that he cannot prevent his eventual loss. However, we observe that player 1 can always choose a vertex $w \in U_2(C)$, since $v \in C = U_1(U_2(C))$; i.e. since $v \in R_1(U_2(C))$. Thus, whenever player 2 picks a vertex in $U_1(B) - C$, player 1 can always respond by choosing a vertex in $U_2(C) - B$ (and his only other choice is to choose a vertex in $R_2(C)$).

Similarly, whenever player 1 chooses a vertex in $U_2(C) - B$, player 2 cannot reach a vertex in C, but he can respond by choosing a vertex in $U_1(B) - C$. So $D_1 = U_2(C) - B$ and $D_2 = U_1(B) - C$: if player i picks a point in D_i, player j ($j \neq i$) can always choose a point in D_j, and must choose such a response to avoid an eventual loss. This completes the proof of the theorem.

As a final note, observe that if player 1, say, picks a winning vertex v in B, then he can assure a win after making m more choices, where m is the number such that $v \in A_{m+1} - A_m$. If player 1 chooses a losing vertex $v \in R_2(C)$, then he can count on making only p moves before losing, where p is the number such that $v \in R_2(C_{p+1}) - R_2(C_p)$.

For example, consider the graph with vertices $V = \{a, b, c, d, e, f\}$ and arcs $A_1 = A' \cup \{(d, e)(e, e)\}$ and $A_2 = A' \cup \{(c, f), (f, f)\}$, where $A' = \{(a, d), (d, c), (c, b), (b, a)\}$. Then it is straightforward to verify that $W_1 = B_2 = \{a, e\}$, $W_2 = C_1 = \{f\}$, $L_1 = \{c, f\}$, $L_2 = \{b, d, e\}$, $D_1 = \{b, d\}$, and $D_2 = \{a, c\}$. Thus, for instance, if player 1 moves to vertex b he can assure himself of a draw, but if player 2 moved to b, he could not prevent a loss.

Of course, in a symmetric game this could not occur: a vertex which is losing for one player would also be losing for the other. For a discussion of the symmetric case, and its relationship to classical concepts of stability in games and graphs, see Roth [2].

REFERENCES

[1] Conway, J. H., *On Numbers and Games*, Academic Press, New York, (1976).

[2] Roth, Alvin E., "Two Person Games on Graphs," mimeo, (1975). Also Journal of Combinatorial Theory, B, *24*, 2, (1978).

[3] Steinhaus, H., "Definitions for a Theory of Games and Pursuit," Myśl. Akademicka, Lwów, *1*, 13-14 (1925). Reprinted in Naval Research logistics Quarterly, *7*, 105-108 (1960).

[4] Smith, C. A. B., "Graphs and Composite Games" Journal of Combinatorial Theory, *1*, (1) (1966).

A NOTE ON EOQ UNDER FUND CONSTRAINTS

Alan J. Kaplan

DRC Inventory Research Office
US Army Logistics Management Center
Philadelphia, Pennsylvania

ABSTRACT

An algorithm is developed to modify the Wilson Q to account for a short-term expenditure constraint over a catalog of items. Representative results are shown and generalizations made.

I. INTRODUCTION

Organizations managing inventories under budgeting systems of limited flexibility periodically face short-term fund constraints. They have a fixed amount of dollars to operate with through the remainder of the fiscal year, at which time they can resume unconstrained operation. The problem arises of how to manage the system with least cost, subject to such an expenditure constraint.

We solve the problem of determining order quantities by employing dynamic programming and a generalized Lagrangian concept sometimes utilized with that technique (cf., Chapter 2, Section 16, of Bellman and Dreyfus [1]). In our case, a Lagrangian cost (Lambda) is defined for each dollar spent before the end of the year. The dynamic programming algorithm then determines, for each item in the catalog which must be bought, how much to cut that item's Wilson EOQ as a function of Lambda, and the time remaining in the year. As the value of Lambda input is raised, solutions will be computed which spend less money.

We do not address in detail the problem of finding the Lambda value which will exactly equate projected funds spent to funds available. Many sophisticated inventory managers, such as the U.S. Dept. of Defense, already have computer programs for projecting total fund expenditures as a function of inventory management parameters, and such programs could be adapted to the purpose of finding a suitable Lambda. In less sophisticated systems expenditures can be monitored, and Lambda used as kind of a control knob. At an even simpler level, the results shown in this paper, obtained from use of the dynamic programming algorithm, indicate which of two alternative "naive" policies for cutting order quantities to save money makes most sense.

II. MODEL

The assumptions and cost structure (apart from Lambda) underlying the Wilson EOQ are assumed. Thus, the costs considered are a fixed cost for each order placed and a holding cost per dollar of inventory held per year. An order for an item must be placed whenever assets

reach 0, and there is no lag placement and receipt of an order. In the classical Wilson formulation demand is deterministic, but following Snyder's [4] generalization we permit demand to be stochastic, requiring only that the interarrival times between successive demands be independent and identically distributed.

For any item purchased there is a dollar outlay equal to the (unit price of the item) × (amount purchased). We wish to minimize total fixed order costs and holding costs over a catalog of items, subject to a constraint on total outlays during the remainder of the fiscal year. By use of the generalized Lagrangian concept, our problem can be solved as a succession of unconstrained single item problems (for each item in the catalog), where we wish to minimize the sum of holding costs plus fixed costs plus outlay costs, costed out at the rate of Lambda per dollar of outlay.

III. BASIS FOR DYNAMIC PROGRAMMING

The need for dynamic programming to solve what has been reduced to a single item problem arises from the limited duration of the constraint. Suppose, for a minute, that demand was deterministic, that the Wilson EOQ for an item was 10, that current assets were 0, and that total demand before year's end was 11. It might well be optimal to make one purchase of 11 units, rather than two smaller purchases; i.e., the impact of the outlay constraint could be to increase the size of the current purchase in order to decrease total outlays on all purchases through year's end in a cost-effective manner. In general, the optimal size of the current purchase cannot be computed independently of its impact on the need for successive purchases.

In order to use dynamic programming we artificially divide the remaining time in the fiscal year into successive review periods. It then is assumed that in any given period we will need to buy at most once, and then only at the beginning of the period. The review period concept is a solution device only, and does not correspond to any entity in our basic model. Why then is its use valid? Well, suppose each review period were only one day long. Clearly, there would be little practical difference between ordering only at the beginning of review periods, or ordering whenever assets hit zero; i.e., we can always get more precise answers by making the review periods smaller. In our use of the algorithm, we used a review period length of 1 week, and then verified that changing the length to 1/4 week had little effect on our answers.

IV. NOTATION AND PRELIMINARY DERIVATIONS

λ — Lagrangian cost per dollar spent before year's end
K — administrative cost to procure
H — cost to hold per year as a percent of unit price
C — unit price of an item
d — expected demand rate (per year)
$G_i(x)$ — probability that total demand over i periods $\geqslant x$
Q_w — Wilse EOQ
$Q_i^*(\lambda)$ — optimum amount to buy with i periods to go

$$S(Q) = [K + (Q/2)\ (Q/d)\ HC]/Q; \tag{1}$$

i.e., S is the expected value of the average holding and fixed order cost incurred per unit when an amount Q is bought. Note that "d" is the reciprocal of the mean interarrival time [3], so $Q\ (1/d)$ is the expected time to deplete Q units and $Q/2$ is the average inventory during that time.

$$P(Q) = [S(Q) - S(Q_w)]\ (Q); \tag{2}$$

i.e., $P(Q)$ is the increase in holding and administrative cost attributable to a buy other than the Wilson EOQ.

$$C(Q,\ \lambda) = P(Q) + CQ\lambda; \tag{3}$$

i.e., $C(Q,\ \lambda)$ is the total cost increase associated with Q, including Lambda cost.

V. DYNAMIC PROGRAMMING FORMULATION

Let $TC(Q,\ i,\ \lambda)$ be the total increase in cost attributable to all buys made during the rest of the year. This assumes there are i periods to go, a buy of Q units must be made now (assets = 0), and an optimum policy is followed thereafter. In other words, $TC(Q,\ i,\ \lambda)$ is the generalization of $C(Q,\ \lambda)$ (equation 3) to include future buys as well as the current buy.

Then,

$$\begin{aligned}
TC(Q,\ 1,\ \lambda) &= C(Q,\ \lambda) \\
TC(Q,\ 2,\ \lambda) &= C(Q,\ \lambda) + G_1(Q)\ TC(Q_1^*,\ 1,\ \lambda) \\
TC(Q,\ 3,\ \lambda) &= C(Q,\ \lambda) + G_1\ (Q)\, TC(Q_2^*,\ 2,\ \lambda) \\
&\quad + [G_2\ (Q) - G_1\ (Q)]\ TC(Q_1^*,\ 1,\ \lambda) \\
TC(Q,\ n,\ \lambda) &= C(Q,\ \lambda) + \sum_{i=1}^{n-1} TC(Q_{n-i}^*,\ n-i,\ \lambda)\ [G_i\ (Q) - G_{i-1}\ (Q)]
\end{aligned} \tag{4}$$

For example, let us look at the equation for $TC(Q,\ 2,\ \lambda)$. First of all, we incur a cost of $C(Q,\ \lambda)$, since by definition of $TC(Q,\ 2,\ \lambda)$ a buy is needed at the beginning of period 2. (If on an item no buy is needed, i.e., assets are above 0, we need not calculate any costs since there is no decision — the only decision we are concerned with is how much to buy when assets hit 0.) In addition to a cost of $C(Q,\ \lambda)$, we will incur an additional cost if another buy is needed at the beginning of period 1. For this to happen, demand in period 2 must exceed what we just bought, (Q), which happens with probability $G_1\ (Q)$. The cost incurred will be $TC(Q_1^*,\ 1,\ \lambda)$. In the equation for $TC(Q,\ 3,\ \lambda)$, $G_1\ (Q)$ is the probability that the next buy will be in period 2, $G_2\ (Q) - G_1\ (Q)$ is the probability that the next buy will be in period 1.

ALGORITHM. The equations are solved recursively. We first find $Q_1^*\ (\lambda)$, the value of Q which minimizes $TC(Q_1\ 1,\ \lambda)$, given by line 1 of (4). Substituting the value found into that equation gives us, by definition, $TC(Q_1^*,\ 1,\ \lambda)$. Next, we solve for $Q_2^*(\lambda)$ and $TC(Q_2^*,\ 2,\ \lambda)$, and so on. To find $Q_i^*\ (\lambda)$ at any step, we have an equation (the appropriate line of (4)), and we just find the value of Q which minimizes the equation. To find the minimizing value we use a grid search, because $TC(Q,\ i,\ \lambda)$ need not be a convex function of Q. In conducting the grid search we look at values of Q in the internal $[1,\ \sqrt{2}\ Q_w]$. In the Appendix it is shown that an optimum Q can never be $> \sqrt{2}\ Q_w$. An optimum Q can be $> Q_w$: although the Lambda concept guarantees that on an expected value basis total purchases over the year will go down, it does not guarantee that all $Q_i^*\ (\lambda)$ must be $\leqslant Q_w$. Recall the example cited in Section III.

VI. RESULTS

Figure 1 gives some representative results. It shows the optimum cut as computed by the algorithm as a function of Lambda and the number of months remaining in the budget year.

CASE	Q_W (IN MONTHS OF SUPPLY)	VARIABILITY	MONTHS TO GO	LAMBDA = 5% 1	2	3	6	LAMBDA = 15% 1	2	3	6
1a	3	Hi		55%	35%	25%	15%	85%	75%	65%	55%
1b	6	Hi		40%	30%	25%	5%	75%	55%	45%	20%
1c	12	Hi		20%	20%	20%	15%	60%	55%	45%	30%
2a	3	Lo		55%	25%	0	0	70%	40%	15%	15%
2b	6	Lo		40%	40%	30%	(10%)	75%	60%	40%	(5)%
2c	12	Lo		20%	20%	20%	20%	60%	60%	60%	35%

FIGURE 1. Cuts in Wilson Q.

For the cases shown, holding cost was at a 25% rate and set-up cost (K) was \$100. A demand rate of 100 per year was used and unit price adjusted as necessary to get the values shown in Figure 1 for EOQ months of supply. High variance denotes a monthly coefficient of variation of 100% while low variance denotes 25%. The negative binomial was used to approximate the demand distribution, as this had previously been found effective for our catalog of items.

Each month had four weeks, the week being the length of the review period. The grid search with intervals equal to (0.05) (Q_w) was used. As a check, cases 1b and 2b were rerun with review period of 1/4 week. The answers never differed by more than (0.05) (Q_w).

VII. CONCLUSIONS

In the high-variance case results were relatively well-behaved. They indicate that a "naive" policy of reducing all buys by the same percent makes more sense than the naive policy of putting some limit on the maximum months in the EOQ as a budget measure. In fact, in the optimum it is the high-dollar items with short EOQ's that tend to be cut the most.

In the low-variance case—one we consider less realistic—the optimum policy may in some situations consist of trying to make one buy to last the rest of the year. Hence, an increase over the Wilson Q may be called for as denoted by the percents in parentheses. This effect can produce a very irregular pattern as shown by these additional results for case (2b):

week	4	8	12	13	14	15	16	20	24
cut	55%	25%	0	(5%)	(10%)	(15%)	25%	15%	0

It is reasonable to implement the algorithm by tables such as those shown. Holding-cost rates and administrative costs are typically constant over at least a major subcatalogue of items. Demand is not, but the only effect of demand, given the other variables specified, is via its effect on the shape of the probability distribution. Thus, for the normal, the results would actually be independent of demand, given the other variables specified, namely: K, H, (Q_w/d), and the coefficient of variation.

For the normal distribution, if you, say, double expected demand (d), but halve unit price and increase variance fourfold (so Q_w/d and the coefficient of variation are unchanged), then Q_1^*/Q_w will stay the same. The normal produces this effect because under the normal density the probability of a value x depends only on the number of standard deviations x is from the mean, not on the mean itself. Hence $G_1(Q) = G_1'(2Q)$, where $G_1(\ )$ assumes an expected demand rate d and standard deviation 1, and G_1' assumes an expected demand rate 2d, and standard deviation 2.

VIII. EXTENSIONS

In principle, the same type of approach can be used to determine cuts in reorder points. The mathematics is such as to require use of approximations. The impact of these is not known, but they are reasonable. To incorporate the proposed cuts in Q_w into the reorder point calculations, however, further complicates the task.

This paper was concerned with a constraint on outlays of a shorter term nature. We assumed that at year's end we could go back to operating in an optimum manner without regard to the pattern of cash outlays. If this is not true, if there is a long-term constraint, it makes sense to react by constraining the total inventory investment. A good paper relating in part to this is that by Presutti and Trepp [2].

While the problem posed in this paper has not, to our knowledge, been treated before, a somewhat related problem has received attention in the production research literature. In that problem, economic lot sizes (read economic order quantities) for a group of items must be determined under a constraint on total production in each period, rather than over a total horizon as here. See, for example, P.S. Eisenhut's paper in AIIE Transactions, June 1975.

APPENDIX

We show that for $Q > \sqrt{2}\,Q_w$, $S(Q) > S(Q/2)$. Therefore, a policy of buying $Q > \sqrt{2}\,Q_w$ would always be dominated by buying $Q/2$, and then making another buy of $Q/2$ later in the budget year as necessary.

$$S(Q) - S(Q/2) = K[1/Q - 2/Q] + (HC/d)(Q/2 - Q/4) =$$

$$(5) \qquad -K/Q + (HC/d)(Q/4)$$

$$(6) \qquad S(Q) - S(Q/2) > 0 <=> -K/Q + (HC/d)(Q/4) > 0.$$

Substituting $HC/d = (1/Q_w^2)(2K)$ in (6), we have

$$(7) \qquad S(Q) - S(Q/2) > 0 <=> -K/Q + \frac{2KQ}{4Q_w^2} > 0$$

$$<=> \sqrt{2}Q_w < Q.$$

BIBLIOGRAPHY

[1] Bellman, R., and S. Dreyfus, *Applied Dynamic Programming* (Princeton University Press, Princeton, N.J., 1962).

[2] Presutti, V., and R. Trepp, "More Ado About Economic Order Quantities (EOQ)," Naval Research Logistics Quarterly *17*, 243-252 (June 1970).

[3] Ross, Sheldon, *Applied Probability Models with Optimization Applications* (Holden-Day, San Francisco, Calif., 1970).

[4] Snyder, Ralph, "The Classical Economic Order Quantity Formula," Operational Research Quarterly *24*, 125-128 (March 1973).

CORRIGENDA

It is regretted that two typographical errors occurred in "Optimal Investment, Pricing and Replacement of Computer Resources," by Charles H. Kriebel, Anthony A. Atkinson and Huntley W. H. Zia which appeared in the December 1977 issue, Vol. 24, No. 4. The errata are as follows:

1. On page 540 the last line of the second footnote was omitted. The footnote should read:

†For convenience and to simplify notation, we will employ the following conventions in the remainder: all functions are assumed differentiable, and, as in (1) and (2), will be designated by capital letters; their derivatives will be written as

$$\partial G_m(a_n,\ b_n,\ c_r)/\partial b_n \equiv g^n_m(b,\ \cdot) \text{ or simply } g^b_m.$$

2. On page 544 the *proof* to *Corollary* (Theorem 2) should read:

PROOF: Trivial from Theorem 2: $MC^+_{tsv}=MC_{tsv^\circ}=MC^+_{ts}\ \forall v$. Q.E.D.

☆ U. S. GOVERNMENT PRINTING OFFICE : 1978 261-252/x-3

INFORMATION FOR CONTRIBUTORS

The NAVAL RESEARCH LOGISTICS QUARTERLY is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Manuscripts and other items for publication should be sent to The Managing Editor, NAVAL RESEARCH LOGISTICS QUARTERLY, Office of Naval Research, Arlington, Va. 22217. Each manuscript which is considered to be suitable material for the QUARTERLY is sent to one or more referees.

Manuscripts submitted for publication should be typewritten, double-spaced, and the author should retain a copy. Refereeing may be expedited if an extra copy of the manuscript is submitted with the original.

A short abstract (not over 400 words) should accompany each manuscript. This will appear at the head of the published paper in the QUARTERLY.

There is no authorization for compensation to authors for papers which have been accepted for publication. Authors will receive 250 reprints of their published papers.

Readers are invited to submit to the Managing Editor items of general interest in the field of logistics, for possible publication in the NEWS AND MEMORANDA or NOTES sections of the QUARTERLY.

NAVAL RESEARCH
LOGISTICS
QUARTERLY

JUNE 1978
VOL. 25, NO. 2
NAVSO P-1278

CONTENTS

ARTICLES		Page
Determining Confidence Bounds for Highly Reliable Coherent Systems Based on a Paucity of Component Failures	J. M. MYHRE A. M. ROSENFIELD S. C. SAUNDERS	213
Approximations to the Reliability of Phased Missions	H. ZIEHMS	229
Inspection Policies for Deteriorating Equipment Characterized by N Quality Levels	Z. KANDER	243
The Error in the Normal Approximation to the Multinomial With an Increasing Number of Classes	L. WEISS	257
A Recursive Algorithm for a Summed Multinomial Density Function	R. K. FINK H. MOSKOWITZ	263
A Probability Model for Initial Crack Size and Fatigue Life of Gun Barrels	F. PROSCHAN J. SETHURAMAN	273
Multiple-Attribute Decision Making With Partial Information: The Comparative Hypervolume Criterion	J. R. CHARNETSKI R. M. SOLAND	279
An Algorithm for a Class of Loading Problems	M. S. HUNG J. R. BROWN	289
A Heuristic Network Procedure for the Assembly Line Balancing Problem	P. A. PINTO D. G. DANNENBRING B. M. KHUMAWALA	299
Pseudo-Monotonic Interval Programming	C. R. BECTOR S. K. BHATT	309
A Targeting Model that Minimizes Collateral Damage	J. H. GROTTE	315
Differential-Game Examination of Optimal Time-Sequential Fire-Support Strategies	J. G. TAYLOR	323
On the Manipulation of Transfer Prices in a Static Environment	J. P. BONIN	357
A Note Concerning Asymmetric Games on Graphs	A. E. ROTH	365
A Note on EOQ Under Fund Constraints	A. J. KAPLAN	369
Corrigenda		375

OFFICE OF NAVAL RESEARCH
Arlington, Va. 22217